

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2002, 02:57:10 : Search time 3627.71 Seconds
(without alignments)
13334.331 Million cell updates/sec

Title: US-09-647-067-7

Perfect score: 3584
Sequence: 1 ctgcttgcactgcctgaaga.....tagnttgcctgaaaccc 3584

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_iny:*
15: em_gss_pln:*
16: em_gss_vrc:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	455.4	12.7	628	9	BB665906	BB665906 BB665906
2	433.8	12.1	469	9	AI089319	AI089319 qbl6d05.x
3	420.4	11.7	444	9	AI805323	AI805323 te32h05.x
4	416.6	11.6	606	10	BG611610	BG611610 602613533
5	405	11.3	454	9	AA122079	AA122079 zk93g10.s
6	404.2	11.3	461	9	AA495807	AA495807 zw05c04.s
7	369.8	10.3	413	9	AI682875	AI682875 wc68d11.x
8	335.6	9.4	567	10	BF181498	BF181498 239107_MA
9	279.4	7.8	544	12	AQ595136	AQ595136 HS_5424_B
10	275	7.7	286	9	AI375172	AI375172 tc10b12.x
11	261.4	7.3	736	9	BB635922	BB635922 BB635922
12	244	6.8	244	9	AA495745	AA495745 zw04c04.r
13	240.4	6.7	292	9	AA122078	AA122078 zk93g10.r
14	235.8	6.6	786	12	AQ745959	AQ745959 HS_2275_A
15	230.6	6.4	593	12	AQ053279	AQ053279 CTR-HSP-2
16	182.2	5.1	227	9	AA626188	AA626188 zv09a01.s
17	175.2	4.9	609	9	AV592706	AV592706 AV592706

c	18	155.4	4.3	744	10	BG304121	BG304121 f130f05.x
c	19	148.8	4.2	702	9	BB297224	BB297224 BB297224
c	20	148.6	4.1	461	9	AV592705	AV592705 AV592705
c	21	144.8	4.0	300	9	AM307854	AM307854 1346 MARC
c	22	128.8	3.6	627	10	BM425549	BM425549 pgf2c.pk0
c	23	124.4	3.5	279	10	BF040869	BF040869 BP250001A
c	24	113.8	3.2	549	12	AQ394779	AQ394779 CIRBI-EL-
c	25	106.4	3.0	235	9	AM436170	AM436170 75600 MAR
c	26	102.2	2.9	532	12	A2319928	A2319928 1M0039D13
c	27	99	2.8	408	12	AQ212353	AQ212353 HS_3002_B
c	28	90.8	2.5	197	10	BF942735	BF942735 EST-CD15N
c	29	87.2	2.4	492	12	A2151363	A2151363 SP_0033_A
c	30	78.6	2.2	516	10	BF158267	BF158267 f130f05.Y
c	31	76.8	2.1	424	12	AQ412148	AQ412148 RPT-11-1
c	32	74.6	2.1	770	12	CNS010W3	AL168348 Tetradon
c	33	74.4	2.1	527	9	BB648375	BB648375 BB648375
c	34	70	2.0	1101	12	CNS05T2L	AL352614 Tetradon
c	35	69.4	1.9	648	9	BB626659	BB626659 BB626659
c	36	68.4	1.9	1101	12	CNS0039C	AL063921 Drosophila
c	37	68	1.8	734	12	CNS010MP	AL099163 Drosophila
c	38	63.8	1.7	1101	12	CNS0039C	AL063921 Drosophila
c	39	62.6	1.7	726	9	AM133623	AM133623 f109f02.Y
c	40	61.8	1.7	2294	11	AK005049	AK005049 Mus muscu
c	41	60.8	1.7	610	10	BI473785	BI473785 f46f08.Y
c	42	60	1.7	582	10	BI122837	BI122837 BI122837
c	43	59.4	1.7	378	10	BE252886	BE252886 601110151
c	44	59.2	1.6	1169	12	CNS06KHQ	AL402900 T3 end of
c	45	59	1.6				

ALIGNMENTS

RESULT	1
BB665906	
LOCUS	BB665906
DEFINITION	BB665906 RIKEN full-length enriched, 2 days pregnant adult female
VERSION	BB665906
KEYWORDS	BB665906.1 GI:1649539
SOURCE	EST.
ORGANISM	house mouse.
REFERENCE	Mus musculus
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 628)
COMMENT	Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Kono,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,H., Sakai,Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shingawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y. RIKEN Mouse ESTs (Arakawa,T., et al. 2001) Unpublished (2001) Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 220-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/ Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) wagl,K., Fujiwara,S., Inoue,K., Togawa,Y., Iwama,M., Ohara,E., Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.

Db 409 AATGTATGGTGTGTGAATAATCTCAGTAAAGCAGTTAAAAAGAAAGAGCTGGAATGC 350
Qy 3261 actgattcagaacttaattcagaaggaaggctgtatgtacacattcatttaag 3320
|||||
Db 349 ACTGATTCAGGAACTTAAATTCAGGAAGAAAGCTGTATGTACACATTTTCACCTTTAG 290
Qy 3321 cagaanaactctctcagaagaatgaacttacttctctcttgcacgtccgaacgtgaat 3380
|||||
Db 289 CAGAAAACTTTTCTTCAAGAAATGACTTTACTTCTCTTTCACATGCGCAGCAGCTGAGAT 230
Qy 3381 actaacttttaactagctgtctctctcgtcgtcgtcgtatgaatltttgtcttc 3440
|||||
Db 229 ACTAACTTTTAAGTACTGTCTTCTCTAGTCTCTAGTATATAGAAATTTTGTCTTC 170
Qy 3441 ataatggaacaccttaagcaggaagaagaatgttttcagatagtttcaatacncaa 3500
|||||
Db 169 ATATATGGAACCTTTTAAAGCAGGAGAAAGAAATGTTTCAATAGTTTCAATACACCA 110
Qy 3501 aaatgttcgacacaaataactcgaatcnaaccataatgcccttattgaatatagt 3560
|||||
Db 109 AAATGTTGAAACACAAATACTGGAATCAACCATATATGCACTTATGAAATATATAGT 50
Qy 3561 tgatagnttctgtctcgaaga 3581
|||||
Db 49 TGTATGATTTGTTCTGAAAA 29

RESULT 3 444 bp mRNA linear EST 13-DEC-1999
AI805323/c
LOCUS
DEFINITION
AI805323
IMAGE:2087673 3', mRNA sequence.
VERSION
AI805323.1 GI:5391889
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 444)
AUTHORS
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 612 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 437.
Location/Qualifiers

FEATURES
Source
1. .444
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2087673"
/clone_1b="Soares_NFL_T_GBC_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: p7T3D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not 1; Site_2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NbHL19W, testis NHT, and B-cell
NCI CGAP GCBI) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."

BASE COUNT 140 a 79 c 60 g 165 t

Query Match 11.7%; Score 420.4; DB 9; Length 444;
Best Local Similarity 99.3%; Pred. No. 2.5e-72;
Matches 443; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

Qy 2319 gagaaactgtgggggtgctcactagaggaattactgtgatgaagaatgaacacaaa 2378
|||||
Db 444 GAGAAATCTGTGGGGGCTTCATGAGGAGATTCTGCTATG-AAATGAATACCAAAA 386
Qy 2379 ttaattataataatagctaaagaataatatttcaagagacatgaggaataaataatg 2438
|||||
Db 385 TTAATTTAATAATATAGCTAAGATAAATATTTTCAAGAGACATGAGGAAAAATG 326
Qy 2439 actaatgctcttaaaaagggaagaatattatacaaatatgatatattatgtgaact 2498
|||||
Db 325 ACTAATGCTCTTACAAAGGGAAGTAATATATATATATATATATATATATATATAT 266
Qy 2499 ttgcataagaataatgaagaatcactcactcagtaacattcattcttctaacatgca 2558
|||||
Db 265 TTG-ATAAGAAATTAAGAAATCTACTCTGTAACATTCATTCTTTCTAATGCA 207
Qy 2559 ttatgagtaaccactactatgtgcataagcattgcaataatagtcctggaagtgaagt 2618
|||||
Db 206 TTTATTTGATACCCACTACTATGTGCATAGCATATGCAATATATAGCTCGAAGTAGACAGT 147
Qy 2619 gcaaaccttcaactcgtatagatagtgttatgaacaaagaacatacaagctcactg 2678
|||||
Db 146 GCAGAACCTTTCATCTGTATAGTGTATATATATATATATATATATATATATATATAT 87
Qy 2679 cagttccctagtttaagtagagcttaccctgtcagtgatgcacgaagaatcagagcac 2738
|||||
Db 86 CAGTTCCAGTTTAAAGTAGAGCTTTACCTGTCAATGTGCATCAGCAAGAAATCATAGGAC 27
Qy 2739 tttaataataaagtttaagtttgg 2764
|||||
Db 26 TTTTAAATAAAGGTTTAAAGTTTGG 1

RESULT 4 606 bp mRNA linear EST 18-APR-2001
BG611610
LOCUS
DEFINITION
602613533F1 NIH_MGC_60 Homo sapiens cDNA clone IMAGE:4738833 5',
mRNA sequence.
ACCESSION
BG611610
VERSION
BG611610.1 GI:13662981
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 606)
AUTHORS
NIH-MGC http://mgc.ncbi.nlm.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
plate: LCM1609 row: a column: 10
High quality sequence stop: 602.
Location/Qualifiers

FEATURES
Source
1. .606
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4738833"
/clone_1b="NIH_MGC_60"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (T1 phage-resistant)"

Db 33 TGTCTGTGACA 24

RESULT 6
AA495807/c
LOCUS
DEFINITION AA495807 461 bp mRNA linear EST 11-AUG-1997
2705C04.s1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:768390
3', mRNA sequence.
ACCESSION AA495807.1 GI:2229128
VERSION
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 461)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Giesel, G., Jost, S.,
Kucab, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B.,
Schellenberg, K., Steptoe, M., Tan, F., Theisling, B., White, Y., Wyllie,
J., Waterston, R. and Wilson, R.
WashU-Merck EST Project 1997
TITLE
JOURNAL
COMMENT Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAG Consortium (info@image.lnl.gov) for further information.
Seq primer: -40m3 fwd. ET from Amersham
High quality sequence stop: 404.
Location/Qualifiers
FEATURES
Source
1. 461
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:768390"
/clone_lib="Soares_NhHMPu_S1"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: pT7T3D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not 1;
Site_2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NbH, pregnant uterus
NbHPU, and fetal heart NbH19W) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."

BASE COUNT 155 a 80 c 67 g 159 t

ORIGIN

Query Match 11.3%; Score 404.2; DB 9; Length 461;
Best Local Similarity 97.9%; Pred. No. 3.7e-69;
Matches 428; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

QY 3147 gtgctgttttaagaagaactaagtgttcaatccaccacttag-atggttgatgt 3205
|||||
Db 461 GTGCTGTTTAAGAAAGACCTTAAGTGTTRATTCACCCACTTTAAGATGGGTAAATGT 402
|||||

QY 3206 tatgtgtgtgaatcatctcgttaagaagctaaagaagaagctgaatgcactga 3265
|||||

Db 401 TATGGTGTGTAATAATCTCAGTAAAGCAAGTTAAAGAAAAAGAGCGAATGCACGA 342
|||||

QY 3266 ttcaagaaacttaattcagaagaagaagctgtatgtacacattcacttaagaacaga 3325
|||||

Db 341 TTCAGGAACCTTAATTTTCAGGAAGAAAGCTGTGTATGTACACATTTCACTTTAAGCAGAA 282
|||||

QY 3326 aatcttctcaagaagaacttacttctcttgcagctgccagcagctgagactaa 3385
|||||

Db 281 AATCTTCTTCAAGAAAGACTTACTTCTCTTGTGCACTGACAGAGTGATCACTAA 222
|||||

QY 3386 cttttaactagtt-gttctctcctagctctcctagcttaagnatttttgccttcata 3444
|||||

Db 221 CTTTCTTACTAGTGTGTTCTCTCTCTAGCTCTAGCTTTTGAATTTTGTGTTCAATA 162
|||||

QY 3445 tctgaacctttaagcagaagaagaatgttttcagatagtttcaaatcncaaaaat 3504
|||||

Db 161 TGTGAACCTTTAAGCAGAGAGAAAGAAATGTTTCAGATAGTTCGAATATACACAAAAT 102
|||||

QY 3505 gtttgcacacaaaatctcgtgaatcnaaccataatgcccattatgatatagtta 3564
|||||

Db 101 GTTGAACACAAAATCTGGAATCAACCATATGACCTTATGAAATATATAGTTGA 42
|||||

QY 3565 tagnttgcctgtaaa 3581
|||||

Db 41 TAGATTGTCTCGAATA 25
|||||

RESULT 7
A1682875/c
LOCUS
DEFINITION A1682875 413 bp mRNA linear EST 17-DEC-1999
wc69d11.x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2323893 3',
mRNA sequence.
ACCESSION A1682875
VERSION A1682875.1 GI:4893057
KEYWORDS
SOURCE EST.
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 413)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE
JOURNAL
COMMENT National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Life Technologies catalog #: 11548-013
DNA sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www-bio.lnl.gov/bbrp/image/image.html
Insert Length: 2559 Std Error: 0.00
Seq primer: -400p from Gibco
High quality sequence stop: 410.
Location/Qualifiers
FEATURES
Source
1. 413
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2323893"
/clone_lib="NCI_CGAP_Pan1"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.72 kb. Life Technologies catalog #:
11548-013"

BASE COUNT 140 a 70 c 58 g 145 t

ORIGIN

Query Match 10.3%; Score 369.8; DB 9; Length 413;
Best Local Similarity 98.2%; Pred. No. 2.1e-62;
Matches 382; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 3193 gatgggtgaatgtatgtgtgtgaatatctcagtaagaagcgttaaaaggaagaagagc 3252
|||||

Db 413 GATGGGTGAATGTTATGTGGTGTGAATATCTCAATTAAGCAGTTAAAGAAAAAGAGC 354
|||||

QY 3253 tggatgcacgtatcaggaacttaattcaggaaggaagagtcgtatgtacacatttc 3312
|||||

[illegible]

RESULT	8
BFI91498	
LOCUS	BFI91498 567 bp mRNA linear EST 02-NOV-2000
DEFINITION	Z39107 MARC 2PIG Sus scrofa CDNA 5', mRNA sequence.
ACCESSION	BFI91498
VERSION	BFI91498.1 GI:11074867
KEYWORDS	EST.
SOURCE	pig.
ORGANISM	Sus scrofa

REFERENCE
1 (bases 1 to 567)
Fahrenkrug, S.C., Fieking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,
AUTHORS

TITLE Design and use of two pooled tissue normalized cDNA libraries for EST discovery in swine

JOURNAL
COMMENT
Unpublished (2000)
Contact: Smith_TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4356
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt-trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR mix

PCR PRIMERS
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTATCCAGTCAGCAG
Plate: 66 row: K column: 22
Seq primer: ATTTAGGTGACACTATAC.

FEATURES	location/qualifiers
source	1. .567

BASE COUNT	ORIGIN
157 a	131 c 148 t

Query Match	9.4%	Score 335.6;	DB 10;	Length 567;
Best Local Similarity	78.8%;	Pred. NO. 1e-55;		
Matches 438; Conservative	0;	Mismatches 114;	Indels 4;	Gaps 3;

OY	1	ctctcttctaactgctaaagattgacgagaaagaatgagcaacaacacgtgtagctgtg	60
Db	15	CTGCTTTGTGGCTGCTTAAGATTGCGAGAGAAATATGCAACACACCGCTGAGCTG	73
OY	61	cgattcagaaaccaagaccnaattttgtcacttcattcattaaatcaagttgtcagatagaag	120
Db	74	CGACTCGGAACCAAGACCAATATATTTACTTTTCATTATACAGTTCTTCCAAAGAG	133
OY	121	gaaatgacatctggttctgtctctctcttcacatcttaatttttggaaatattttctcat	180
Db	134	GAAATGAAATCTGGGCGCCGTCCTTCATCTACACTTGATTTTGGAAACTATTTGCTCT	193
OY	181	gggggttgacagagatgtaagatgctcccttgatcttcccttgggggaatcacaaag	240
Db	194	GGGAGTGGGCGAGATGTATACATGCCCCCTGGGCTACTTCCCTGTGGAGACATCACAAg	253
OY	241	tgtctgctcagctcctgtcactgtaaagcgtgtgacagactcgaggaaatcagccgatgag	300
Db	254	TGCTTGCCCTCAGTTGCTTCTACATCGCAACGGTGTGAGCACTGTGGGAACACGGCTACAG	313
OY	301	gacaactgtgtgtgtgttttgtccagtcatg-tctttgccaggtctggaactgtaactg	359
Db	314	GACAACGTGGAGACAAACATGAGTGGCTATGCAACTTGACAAATATTTGGCAATTAC	373
OY	360	gatgaacaccttaagagtggtccatcggtttcttccaatagtaagtaagttaacttca	419
Db	374	TACAATAATGACTCCCTCCATCTACTTTTGAAGCGACAGACATCTGAAATGC--TGTCAC	431
OY	420	gtgagacttaataagaagaagcttcctccctgtaattgtcttcaagaattatcatgacttc	479
Db	432	GTGGAAATTAATAAATAAGCTTCCTCCGATGGGCTTCAAAAAGTACCATGATCTTCAGAA	491
OY	480	gctggaactgcaaaacaataagatataatccatccatccatctatgctttcagaagactgaa	539
Db	492	GCTGTGCTGCAAAACAATTAAGATTAGGTCCGTATCTGCTATAGCTTCCGAGGACTGTA	551
OY	540	tagctacttaactag	555
Db	552	CAGTCTCACTTAAC	567

RESULT	9
A0595136/C	
LOCUS	
DEFINITION	544 bp DNA linear GSS 08-JUN-1999
ACCESSION	A0595136
VERSION	HS-542L.B1.E06.T7A RPCI-11 Human Male BAC Library Homo sapiens
KEYWORDS	genomic clone Plate-1000 Col-11 Row-J, DNA sequence.
SOURCE	A0595136
ORGANISM	A0595136.1 GI:5026722
REFERENCE	GSS.
COMMENTS	human.

ORGANISM
Homo sapiens
Euarchyotia, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 544)
REFERENCE
AUTHORS
Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T., et al.

TITLE	JOURNAL	MEDLINE	COMMENT
Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome	Proc. Natl. Acad. Sci. U. S. A.	96 (17), 9739-9744 (1999)	99380589
Contact: Mahairas GG, Wallace JC, Hood L			

High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dsj.org, med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)

Thu Sep 5 14:39:12 2002

us-09-647-067-7.rst

Page 11

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2002, 02:58:00 : Search time 6304.42 Seconds
(without alignments)
11896.528 Million cell updates/sec

Title: US-09-647-067-7

Perfect score: 3584

Sequence: 1 cgcgttctaactgctaaga.....tagntttgtctgaaacc 3584

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBank: 1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Match Length DB ID Description

C	1	1996.4	55.7	2274	9	AF190500	AF190500 Homo sapi
	2	1571.4	43.8	187089	2	AC107219	AC107219 Homo sapi
	3	1189.4	33.2	1191	6	AX128507	AX128507 Sequence
	4	687.8	19.2	2436	9	AF453828	AF453828 Homo sapi
	5	607.6	17.0	2539	10	AF346501	AF346501 Mus muscu
	6	439.2	12.3	474	6	AX088165	AX088165 Sequence
	7	437.2	12.2	1068	6	AX148176	AX148176 Sequence
C	8	247.8	6.9	171000	2	AC098607	AC098607 Rattus no
	9	246.2	6.9	575	5	AX356204	AX356204 Sequence
	10	194.2	5.4	3756	3	LSGPCR	Z23104 L. stagnalis
	11	187.4	5.2	170522	9	AL138708	AL138708 Human DNA
	12	187.4	5.2	176565	2	AL136106	AL136106 Homo sapi
	13	175.4	4.9	171000	2	AC098607	AC098607 Rattus no
	14	173	4.8	125322	2	AC108017	AC108017 Homo sapi
	15	173	4.8	180122	9	AC019341	AC019341 Homo sapi
C	16	162.8	4.5	180673	10	AC068627	AC068627 Mus Muscu
	17	162.8	4.5	213462	10	AC077689	AC077689 Mus Muscu
	18	158.4	4.4	592	6	AX320420	AX320420 Sequence
C	19	138.2	3.9	125322	2	AC108017	AC108017 Homo sapi
	20	114.2	3.2	576	6	AX320477	AX320477 Sequence
C	21	110	3.1	1018	6	AX147820	AX147820 Sequence
	22	103.8	2.9	155163	2	AC098990	AC098990 Rattus no
	23	96	2.7	1015	6	AX254453	AX254453 Sequence
C	24	89.2	2.5	155163	2	AC098990	AC098990 Rattus no
	25	85.8	2.4	156490	10	AL603709	AL603709 Mouse DNA
C	26	81.6	2.3	67157	2	AC101958	AC101958 Mus muscu
	27	76	2.1	10197	1	AE006607	AE006607 Streptoco
C	28	76	2.1	120871	9	AC016605	AC016605 Homo sapi
	29	73.6	2.1	123050	3	AC008917	AC008917 Homo sapi
	30	73.6	2.1	2895	3	DMTARTAN	U02078 Drosophila
	31	73.6	2.1	3192	3	AY051439	AY051439 Drosophila
	32	73.6	2.1	96877	2	AC017776	AC017776 Drosophila
C	33	73.6	2.1	162343	3	AC010659	AC010659 Drosophila
	34	73.6	2.0	171729	3	AE003539	AE003539 Drosophila
	35	72.2	2.0	1080	6	AX254459	AX254459 Sequence
	36	68.2	1.9	2019	6	AX179624	AX179624 Sequence
	37	64.4	1.8	5261	10	AF144627	AF144627 Mus muscu
	38	64.2	1.8	17443	5	AC024806	AC024806 Caenorhab
	39	63.8	1.8	3075	3	AB005587	AB005587 Cynops py
	40	63.2	1.8	5954	9	AB058757	AB058757 Homo sapi
	41	63.2	1.8	8594	9	HSY19205	Y19205 Homo sapien
	42	63.2	1.8	161341	9	HSBG115M3	AL109653 Human DNA
	43	62.4	1.7	4015	9	AB018349	AB018349 Homo sapi
	44	60.4	1.7	1496	3	CEY5312D	AL033123 Caenorhab
	45	60.4	1.7	2621	9	AB046639	AB046639 Macaca fa

ALIGNMENTS

RESULT	1	AF190500	2274 bp	linear	PRI 10-AUG-2001
LOCUS	AF190500				
DEFINITION	Homo sapiens leucine-rich repeat-containing G protein-coupled				
VERSION	AF190500				
KEYWORDS	AF190500.1 GI:10441729				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 2274)				
AUTHORS	Hsu,S.Y., Kudo,M., Chen,T., Nakabayashi,K., Bhalla,A., van der				
TITLE	The three subfamilies of leucine-rich repeat-containing G				
	protein-coupled receptors (LGR): Identification of LGR6 and LGR7				
	and the signaling mechanism for LGR				
JOURNAL	Mol. Endocrinol. 14 (8), 1257-1271 (2000)				
MEDLINE	20388592				
PUBMED	10935549				
REFERENCE	2 (bases 1 to 2274)				
AUTHORS	Hsu,S.Y. and Hsueh,A.J.W.				

TITLE Direct Submission
JOURNAL Submitted (28-SEP-1999) GYN/OB, Stanford University, MSOB S385, Stanford, CA 94305, USA
FEATURES Location/Qualifiers

```
source      1. . 2274
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
gene        1. . 2274
            /gene="LGR7"
CDS         1. . 2274
            1..2274
```

```

/codon_start=1
/product="leucine-rich repeat-containing G protein-coupled
receptor 7"
/protein_id="AAG17167.1"
/db_xref="GI:10441730"
/translation="MTSGSVFFYLLIFGRKFSHGGGDVRCSLGYPECGNITKCLPDL
LHCGVDDCGNQADEDCGDNNGMSWQFDKFAFYKMTSQYPEAATPECLQGSVSV
OCLOGLGELDCDEFTNLRAVPSVSNVTAQNNMLIRKLPDPCRKANYHDLQKLYQNN
KITSSTIYAFERGLNSIKRLYSNHRITFEFGVEDHRLFWMLIIEENHLSRISEPFEE
YGNLSILLIYNNVLTLPDPKPICOMPRHLMTDLGNNHNLKNTLSFSCSLYLTQKI
VMRNKKNHNLNENTFAPLQKIDELDLSKNRNTPLPTFRPKLRLSOLNTSNPIQKIK
QANQFVLVLRKLSLGIELSTNQQRERFRLMLSHIRFKKPYCGYAPRVRSCKRPR
TDLGISLVENLGLAIQRFVWVWVASVATFCGFIYICMSPYIRSNKNTFMSJLISCCA
DCLMGIVLFIYFGDLKFRGEGYNHAQMLMSTGCQLVSLALISTEVSILLFLLFTL
EKYCIYVPERCVARGKCRITTVLLIWMITGFIYAALFDPSKSEFKKNYGGKCPFL
HSEPTESGIAQIYVALEFGINTLAALFIYVSGSMSEFVSQAATATEINQVCKEPL
ILARFPFIYFTDALCMPIFVVEKFLSLQVEIGTITSWVFIPLPNSALNPLILVTA
LTTPRPEMIIHRFVNYRORKSMDSKGQKTYAPSEIYEMWMPLOEMPELAKPLDFTFY
PCENSLISQSTRINYS"

```

Query Match	55.7%	Score 1996.4	DB 9	Length 2274	
Best Local Similarity	94.9%	Pred. No. 0			
Matches 2158	Conservative	0	Mismatches 11	Indels 105	Gaps 5
QY 124 atgacatcgggtcgcgtctctctcaacatcttaattttggaaaatatttccatggtg				183	
DB 1 atgacatctgggttcgtctctctcttcttccatcttatttttggaaaattttttctcatggg				60	
QY 184 ggtggacaggaatgltcaagtgcctcccttggctatltccccctgtgggaaacatcaaatgtgc				243	
DB 61 ggtggacagagatgtcaagtgctcccttggtgcattttccctgtgggaaatcacaagaatgct				120	
QY 244 ttgcttcagctccctgcactcgttaagggtgtgtggacgactgtggygaatcagcgcatgagac				303	
DB 121 ttggcctcacacttccctgcactgtcaacgtgtggacgactgcgggaatcagggcccatgagac				180	
QY 304 aactcttg-----				310	
DB 181 aactgtggagacaaacatgatgatgtccatgcatttggacaaatattttgccagttaactac				240	
QY 311 -----				323	
DB 241 aaatgacttcccaatattcctttttgagcagaaaacacatgaaattgtttggctgcttctgtgc				300	
QY 324 ccaagtc-atgtcctttgcc-aagctcggagcttgaact-gatgaacaa--tttaagagt				378	
DB 301 ccagtcacatgttctttggccaaaggtggagctgtgactgtgatgttaaaaccaattttacgagct				360	
QY 379 gtccacatgggtttcttccaatgtgaactgaatgtcaacttcagtygaacttaataagaag				438	
DB 361 gtttccatcggtttcttccaatgtgaactgaatgtcaacttcagtygaacttaataagaag				420	
QY 439 ctctctccgtatgtgttccaagaattcatcaltcctcaagccttgagcctgcacaaacaat				498	
DB 421 ctttcctccgtatgtgttccaagaattcatcaltcctcaagccttgagcctgcacaaacaat				480	
QY 499 aagattacatcatcctcatctatgcttccagagagactgaatagcttacttaactgtat				558	
DB 481 aagattacatcctcatcctcatctatgcttccagagagactgaatagcttacttaactgtat				540	

QY	559	ctcagtcataacagataaactctccctgagccgggtgttttgaaagactctcaagacta	618
Db	541	ctcactctatAACAGATTAACCTTCTGTGAAGCGGGGTGTTTTGGAAGATCTTCACAGACTA	600
QY	619	gaatggctgataatggaatataactgaactgaattcccaattcccaacatttatga	678
Db	601	GAATGGCTGATTAATGGAATATATACCTGACGTCAATTTCCCCACCAATTTTATGGA	660
QY	679	ctaaattctctatctctctctgaactgataaagcttccctcaaccggttacttgataa	738
Db	661	CTAAATTCCTTATTCCTCTTACTGATGATTAACGTCACCCGTTTACTGATTA	720
QY	739	ccctctctgcaacacatgccaagactacaatgctgagacctgaaagccaacatctcat	798
Db	721	CCCTCTCTGCAACACATGCGCAGACTACATTTGGCTGGACCTTGAAGCGCAACCATATCAT	780
QY	799	aatttaagaattgaactttatctcccgagtaattaaactggtttagtgtgaggaa	858
Db	781	AATTTAAGAAATTTGACTTTTTCCTGCAATAATTAACTGTTTAACTGATGAGGANA	840
QY	859	aacaaataatcacactaaatgaataatacttttgacccctcccaagaatctgatatg	918
Db	841	AACAAATAATTAACACTTAAATGAATACTTTTGGACCTCTCCAGANAATGGATGATTG	900
QY	919	gatttagaagaataaagattganaaactctccacgcgtatataccaagacctgaagag	978
Db	901	GAFTTAGAAGATTAATTAAGATGAATGAATCTTCCACCGCTTAATTAAGACGCTGAAAGAG	960
QY	979	ctgtccaaattgaactcttccctataatccaaatccgaagaattcaagcaacaaattgct	1038
Db	961	CTGTCACAATTAATTAATCTTCTCTAATATCCAAATCCAAATAATTCAGCAAAATTTGAT	1020
QY	1039	tatctgtgtaaaactcaagctctccagccctagaagagatgaaattccaataccaaca	1098
Db	1021	TATCTGTAAACTCAAGTCTCTCAGCCTCAGAGAAGGATTTGAATTTCAATATTCACANA	1080
QY	1099	aggatgtttagaaccttita tgaactctctccacatalatttaagaattccaagtact	1158
Db	1081	AGGATGTTTAGACCTCTTAATGAATCTCTCTCAATATATTTTAAGAAATTCAGTACTGT	1140
QY	1159	gggatgacccacatgctgcgcagctgtaaaccaaacatgtagaattcatctctcagg	1218
Db	1141	GGGTATGACACACATTTGTGGACCTGTAACCAAAACGATGGAATTTCACTCTAGAG	1200
QY	1219	aatctcttggaagacttattcagagaagatgtgtctggtgtactgcagttactcgc	1278
Db	1201	AATCTTTGGCAAGCAATTTATTCAGAGATTTGTGTGGGTGTATCTGCAGTTACTCTGC	1260
QY	1279	tttggaaacaattttgttcatattgcatgagacattalatacgaagctcgtgagaaacgctgat	1338
Db	1261	TTTGGAAACAATTTTGTGCATTTGCATGCGACCTTATATCAGGCTGTGAGAAACAAGCTGAT	1320
QY	1339	gccaatgcaatcaatctctctcgtcgtgtgcgcagctgttaatgggaatatattatctcgt	1398
Db	1321	GCCATGTCAATCATTTTCTCTCTGCTGTGCGACTGCTTAATGGAATATATTTATTCGTG	1380
QY	1399	atcgagagcttgaactaaagtttctgtgagaatacaataaagatgcgcagctgtgatg	1458
Db	1381	ATCGGAGGCTTTGACCTTAAGGTTTCGTGTGGAATTAACATTAAGCATGCGCAGCTGTGGAG	1440
QY	1459	gaagatcactcatgctcagcttgttaggactcttggccattctgtccacagaagatcaagct	1518
Db	1441	GAAGATACATATGTGCAGCTTGAAGATCTTTGGCCATCTGTGCCACAGAAATATCAGTT	1500
QY	1519	ttactgttaacattctgacattggaataaacatcatctgcatgtctcatccttttaagt	1578
Db	1501	TTACTGTTAACATTTCTGACATTTGAGAAAATATACATCAGCATTTCTTAATCCTTTTAACTGT	1560
QY	1579	gttagagactgtgaaaatgacgaacaataaagctctgataatctcatctcttggaattactg	1638
Db	1561	GTAGAGCTGTGGAANAAGCAACAATATTAACGTTCTGATTTCTCATTTTGGATTAAGTTTACTG	1620

```
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
```

```

Contact: submissions@watson.wustl.edu
Project Information
Center project name: H_NH0575B04
-----
Summary Statistics
Sequencing vector: M3; 0%
Sequencing vector: plasmid; 100%
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 182907 bases at least Q40
Consensus quality: 183613 bases at least Q40
Consensus quality: 184039 bases at least Q20
Insert size: 200000; agarose-fp
Insert size: 186189; sum-of-contigs
Quality coverage: 6.48 in Q20 bases; agarose-fp
Quality coverage: 7.47 in Q20 bases; sum-of-contigs
-----
NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1
* 3402: contig of 3402 bp in length
* 3403
* 3502: gap of unknown length
* 3503
* 10112: contig of 6610 bp in length
*
* 10113
* 10212: gap of unknown length
*
* 10213
* 23531: contig of 13319 bp in length
*
* 23532
* 23631: gap of unknown length
*
* 23632
* 34706: contig of 11075 bp in length
*
* 34707
* 34806: gap of unknown length
*
* 34807
* 53512: contig of 18706 bp in length
*
* 53513
* 53612: gap of unknown length
*
* 53613
* 113083: contig of 59471 bp in length
*
* 113084
* 113183: gap of unknown length
*
* 113184
* 175773: contig of 62590 bp in length
*
* 175774
* 178033: gap of unknown length
*
* 175874
* 178033: contig of 2160 bp in length
*
* 178034
* 178133: gap of unknown length
*
* 178134
* 182702: contig of 4569 bp in length
*
* 182703
* 182802: gap of unknown length
*
* 182803
* 187089: contig of 4287 bp in length.
Location/Qualifiers
FEATURES
source
1.187089
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
/clone="RP11-57584"
1.3402
/misc_feature
/note="assembly_name:Contig10"
3503.10112
/misc_feature
/note="assembly_name:Contig11"
10213.23531
/misc_feature
/note="assembly_name:Contig12"
23632.34706
/misc_feature
/note="assembly_name:Contig13"
34807.53512
/misc_feature
/note="assembly_name:Contig14
vector_side:right"
clone_end:SP6
53613.113083
/misc_feature
/note="assembly_name:Contig15
clone_end:T7
vector_side:right"
113184.175773
/misc_feature
/note="assembly_name:Contig16"
175874.178033
/misc_feature
/note="assembly_name:Contig17"
178134.182702
/misc_feature
/note="assembly_name:Contig18"

```


ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 2539)
 AUTHORS Overbeek, P. A., Gorlov, I. P., Sutherland, R. W., Houston, J. B., Harrison, W. R., Boettger-Tong, H. L., Bishop, C. E. and Agoulnik, A. I.
 TITLE A transgenic insertion causing cryptorchidism in mice
 JOURNAL Genesis 30 (1), 26-35 (2001)
 MEDLINE 21250990
 PUBMED 11353515
 REFERENCE 2 (bases 1 to 2539)
 AUTHORS Agoulnik, A. I.
 TITLE Direct Submission
 JOURNAL Submitted (06-FEB-2001) Ob/Gyn, Baylor College of Medicine, 6550 Fannin Str., Su. 861, Houston, TX 77030, USA
 FEATURES
 source
 1. 2539
 /organism="Mus musculus"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /chromosome="5"
 1. 2539
 /gene="Great"
 /gene="Great"
 /note="Great"
 /codon_start=1
 /product="G protein coupled receptor affecting testicular descent"
 /protein_id="AA108943.1"
 /db_xref="GI:15811373"
 /translation="MMLLHVLLEVEKDFALDSSWVAPLCPKYPGCLNLRKLP
 AFHDGDVDCNGADENDCGDTSGWTPICTGVNKKVTLDECPISQYPOHCYCE
 NEHECYADIKAVPKVSSNTLLSKKKIHRPDKVSETRTEKRTYLDHNCITHS
 RRAFLGHNLIQILYLSHNCITSLRPGLEKQLHQLAMLLIDNPTRISQKSEGLNSL
 FLPWGNRLDEALPETLCAQMPQMLWYDLNANGIKYITNSFTLIDSLTFLVLPKRI
 GPVEKFFSKNLGELDSSNMITKLPVLEFSDLQKLSNLDLYVHKNOFS
 LKQLOSLDERIEIPNISTGMFOPMKNLSHLYKTFPCYGVPHVRINCPSTDGISS
 EDLLANGILRSVWVAFITFCVGNFIVAVRSILKANTHANSIKITLCADCLMGY
 LRSYGVFDIKRYOKYKATLLMESVPCRLIGLATISTEVSYLLTFLLEKLVIV
 FPPSNLRGRQIRVALASTWVGFLLAAVPTREDEYGFNGVCGEPLVADQAD
 FGSRGYSLGFLVNLAFLLVIVSYVMCSLHKFALQTAEVSHGKEVAANRF
 FIVESDAICWIPVFWKILSLQVEIPGITSMTVVFLLPVNSALNPILYLTITSFK
 DKLKQLHKRRKRPFKVKKKSASIVMTDESLKGLVSLKIALGDSIMKPYSP"
 BASE COUNT 639 a 641 c 574 g 685 t
 ORIGIN
 Query Match 17.0%; Score 607.6; DB 10; Length 2539;
 Best Local Similarity 59.8%; Pred. No. 6.5e-114;
 Matches 1037; Conservative 0; Mismatches 694; Indels 3; Gaps 1;

Db 835 GTTAACTCCTTGTCTTTCTTGGCCATGTTGGGTAACCGGTGACGGCCCTCTCTG---A 891
 Qy 738 accctctgtcaacacatgcgaactacatctgctgaccttgaagcaacatcatca 797
 Db 892 AACATTGTCTCAGATGCTCACTCACTCACTGCTGATGCTGCAACATGAAATAA 951
 Qy 798 taattgaattgtcttattcttccgtgagtaatttaattgctttagtgatgagaa 857
 Db 952 GTACATTAACGAACTCACTCTTCTTACGCTGACCTGCTCAGGTTCTGTTGCTCTAG 1011
 Qy 858 aacaaaataactaactaaataactcttgcacccctccagaacatcgtatgaatt 917
 Db 1012 AAATCAATTTGCTTTTGTTCAGAGAAAGACATTTCTTCAATTAATAATTAGAGAACT 1071
 Qy 918 ggattcggagtaataagatgtaaaatcttcacccgttatattcgaagccctgaagga 977
 Db 1072 GGACCTGTCTGCAATATGATGATACAAACCTCCAGTCCACTTCTTCAAGCCATTCATCT 1131
 Qy 978 gctgtcaaatgtaattcttcttcttaataccaatccagaanaattcaagcaacaaattga 1037
 Db 1132 TCTCCAGAGCTGAACCTGTATCCAACTCTTCTGTATGTCACAGAAACGAGTTGG 1191
 Qy 1038 ttatctgtcaactcaagctctcctcgaagctgaagggattgaaattcaaatatccaaca 1097
 Db 1192 AAGTCTCAAACTGATGCTGCTAGACCTGGAAGATGATGATCCAAACATTAAGCAC 1251
 Qy 1098 aagatgtttagaccttatgaatctctcacaatatatttaagaataatccagtaagt 1157
 Db 1252 AGGATGTCTCAGCCAAAGAAACCTTCTCACATTTATTTGAAACCTTTCATATCTG 1311
 Qy 1158 tgggtatccaccatctgtcgaagctgaagctgaacaaacacatgtaatttcatctcaga 1217
 Db 1312 CTCTATGTCCCATGCTCCAACTGTATGCTCCGACATATGATTTTGTGCTCGA 1371
 Qy 1218 gaactctgtgcaagcatattcagagagatattgtctggtgttatctgcagttacgt 1277
 Db 1372 GGACCTCTGGCTAAAGGATATCTCAGAGTGTGCTGCTGGTTATAGTTTCATTACTCG 1431
 Qy 1278 ctgtggaacatttctgcaattgcattgcagagcactatatacagctctgagaacacgtga 1337
 Db 1432 CGTTGGGAATTTCTTGTCAATGACCGGTGATGCTCTCATTAAGCTGGAATTAACATCA 1491
 Qy 1338 tgcacgtcaatcaatctctcctcgtctgctgagcagctcttaatgggaatatattatcgt 1397
 Db 1492 CGCTATGTCATCAAAATCTCTTTGTTGTGCCGATTTGCTGATGGGGGTGTAACCTGTC 1551
 Qy 1398 gacggaagcttgaactaaagttcgtggaatatacaataagcatgagcagctgtgagat 1457
 Db 1552 CGTGGCGCTTTTGTGACATCAATGACCGAGGCGAGATGACGAATGATGCGTGTGAT 1611
 Qy 1458 ggaagatcaatctgctcagctgtgagatcttggcactctgtccacagaagatcagct 1517
 Db 1612 GGAGAGTGTGCCCTGCCCTGCTGGGCTTCTCGCCAGCTGTGCCACAGAGGTCCTCGT 1671
 Qy 1518 ttactgttaacattctgcaattgggaanaataacatcgcgacttctcactctttagatg 1577
 Db 1672 GCTGCTGCTGACATTTCTGACAGCTGAGAAATCTCTTGTCAATGATATTCCTTTACGAA 1731
 Qy 1578 tgtgagaccttgaaatgcaacaacatcagttctgattctcatttggattactggtt 1637
 Db 1732 COTGGCGCTTGGGCAAGGCCAGACTGCTGTGGCCCTGCCAGCATCTGTGGGTGGGAT 1791
 Qy 1638 tatagtgcttcatctcattgagcaataagaatttctcaaaaactactatgaccaa 1697
 Db 1792 TCTCATACCGGCGCTTCCGTTCCACAGAGAGATTAATTTGCGCACTTTTATGGAAAAA 1851
 Qy 1698 tggagatgcttccctctcattgaagaatgaagaagaatttggagccagattatctc 1757
 Db 1852 TGGAGTCTGCTTCCACTCTATATGACCAAGCAAGAAATTTGGAGTAAAGGTTACTC 1911
 Qy 1758 agtggcaatttctctgtatataatttggcgaatttcatcactagatttctcattag 1817

Db 1912 CCTGGATTTCCTAGGTGTAACCTGCTGCTTCCTCGTCATCGATTCCTANGT 1971
Oy 1818 aagcattgtttatagtttcatcaagaatgccaataacagcaactggaataacgaatcaagt 1877
Db 1972 CACCAATGTTCTCTCCATTAATAAACACCCCTTCAAGCTGAGAAAGTAAAGAGCACAT 2031
Oy 1878 taaaaaagagatcattctgccaacagttttctttatagatttacttactgattatg 1937
Db 2032 CGGGAAGAGAGGGGTGTGGCAACCGGTTCTTTTATCGTGTCTCGATGCCATCTG 2091
Oy 1938 ctgataaccatttcttgtagtgaattcttcaactgctcaggtagaataaccagttac 1997
Db 2092 CTGGATCCCTGTTGTTGCTTAAGATCTGTCCTCTCAAGTGAAGATACCAAGCAC 2151
Oy 1998 ctaaacctcttggtgtagtatttattcttgcacatlaacagtgcttgaacccaattct 2057
Db 2152 AATCACTTCCTGAGATCGTGTGTTTTCCTTCGCGGAACAGCGCTTAACCCCATCT 2211
Oy 2058 ctatactctgacacaaagacacattaaagaatgattcatcggtttgtataa 2111
Db 2212 CTACACTCTGAGACCTCCTTTTAAAGACAAGTTGAACAGTGTGTCACACA 2265

RESULT 6
AX088165 474 bp DNA linear PAT 17-MAR-2001
LOCUS AX088165 Sequence 1 from Patent WO0114548.
DEFINITION AX088165
ACCESSION AX088165 GI:13397078
VERSION AX088165.1
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 474)
AUTHORS Duecker,K.
TITLE New g-protein coupled receptor and dna sequences thereof
JOURNAL Patent: WO 0114548-A 1 01-MAR-2001;
MERCK PATENT GmbH (DE)
FEATURES
source 1.474
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
<1..474
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC34686.1"
/db_xref="GI:13397079"
/translation="KOISVALFLGINLAFTIIVPSYMSFVSHQSAITATETIRNO
VKEMILARFEFFIVFDALCWIPLEVARPLSLDOVEIPGITSWVYIYSAINSLN
PIIYLTLPPEKEMTHREWHNYRORSMDSGIRKRLHMHSSGGKGRCHLS"
BASE COUNT 144 a 92 c 88 g 150 t
ORIGIN

Query Match 12.3%; Score 439.2; DB 6; Length 474;
Best Local Similarity 96.8%; Pred. No. 1.7e-79;
Matches 459; Conservative 0; Mismatches 13; Indels 2; Gaps 1;

Oy 1744 gcccaagattatccagtggaattttctgtgattatttgccgcatatcatata 1803
Db 1 GCCCAGATTATTCAGGCGCAATTTTCTGGATTATTAATTTGGCGCATTTATCATATA 60
Oy 1804 gtttttccatggaacatgatttattagtttcatcaaatgccaataacgaactgaa 1863
Db 61 GTTTTTCCTATGGAACATGTTTATAGTTCATCAAAAGTGCATTAACAGCAACGAA 120
Oy 1864 atacgaatcaagttaaaaaagagatgactctgccaacagtttttcttatagtattc 1923
Db 121 ATACGAATCAAGTTAAAAAGAGATGATCTTGCCAAAGCTTTTCTTATAGTATTT 180
Oy 1924 acgatgcatatgctgataccattttgttgtagtaattcttcaactgctcaagta 1983
|||||

Db 181 ACTGATGATTATGCTGCATACCCATTTTGTAGGGAACCCCTTCACTGCTCAGGA 240
Oy 1984 gaataaccaggtaccataaccctcttggttagtgattttatcttgcacattacagttct 2043
Db 241 GAAATACCAAGGTACCATTAACCTCTTGAGTAGATGTTGTTATTTCCATTAACAGTCT 300
Oy 2044 ttgaacccaattcttatactcttgaccacaagacatttaagaatgattatcagttc 2103
Db 301 TTGAACCCAAATTCCTATACCTGACCAACAGACATTTAAAGAAATGATTCATGCGTT 360
Oy 2104 tggtaactacacagaagaataatctatgagacagaag--tcaagaataatgctc 2161
Db 361 TGGCAATTAATACAGCAAAAGAAATCTATAGACACAAAGATATCAGAAATATGCTC 420
Oy 2162 catcattcatctggttggaatgtagcaactgcaagagatgcaccctgattaa 2215
Db 421 CAVCAATTCATCTGGGGGGAATGTGGCCACTCAGAGATGACCACTGAGTTAA 474

RESULT 7
AX148176 1068 bp DNA linear PAT 08-JUN-2001
LOCUS AX148176 Sequence 17 from Patent WO0136471.
DEFINITION AX148176
ACCESSION AX148176
VERSION AX148176.1 GI:14347081
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1068)
AUTHORS Chen,R., Dang,H.T. and Lowitz,K.P.
TITLE Endogenous and non-endogenous versions of human g protein-coupled
JOURNAL receptors
Patent: WO 0136471-A 17 25-MAY-2001;
Arena Pharmaceuticals, Inc. (US)
FEATURES
source 1.1068
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 278 a 215 c 214 g 361 t
ORIGIN

Query Match 12.2%; Score 437.2; DB 6; Length 1068;
Best Local Similarity 65.9%; Pred. No. 4.1e-79;
Matches 650; Conservative 0; Mismatches 333; Indels 3; Gaps 1;

Oy 1195 actgaggaattcatcctcagagaatcctctgcaagcatattcagagagatttgc 1254
Db 10 ACGGACGGCATTTCTTCATTTGAGGACCTTGGCTAACAAATTCCTCAAAATATTGTC 69
Oy 1255 tgggttgatctgcaagttacccgtcttggaacattttgtcatcttgcatgtagacttat 1314
Db 70 TGGGTATATCTTTCATTTACTCTGCTTGGAAATCTTTGTGCAATGGCAGATCTTTC 129
Oy 1315 atcagtcctgagaacaagctgtagcatgcaataattctctctgcttgcgcagctgc 1374
Db 130 ATTAAGCTGAATAATACACTGACGTATGTCATTAATAATCCCTTGTCGCTGATTGC 189
Oy 1375 ttaatggaatatattatcttgcgtatcgtagcggtcttgacctaagttctgtggaataac 1434
Db 190 CTGATGGGTATTACTTGTCTTGTGGCATTTTGCATTAATAATACGAGGAGATAT 249
Oy 1435 aataagcatgcaagctgtagatgagagatcattatgtaactgttagatcttggcc 1494
Db 250 CAGAGATATGCTTGTGATGAGAGCGTGCAGTGCAGCCCTCAAGGGGTTCCCTGCC 309
Oy 1495 attcttccacagaagatcagtttactgtttaacattcttgcaattggaataataatc 1554
Db 310 ATGCTGTCACCGAAGCTGTGTTGCTACTGACCTACTTACTTGGAGAACTTCTG 369
Oy 1555 tgcatgtctatccttttagatgtagttagaactggaataatgcaataatcagttctg 1614
|||||

Db	370	GTCATTTGCTTCCCTTCAGTAACATTTGACCTCGAAGAAAAGGCACAGCCATCATCTCC	429
QY	1615	attctcaattggatactacggtttatagtgcttcaatccatlyagcaataagaatt	1674
Db	430	ATTTCACATCGATGCGGGGATTTTAAATAGCTGAATTCCTTTGGAAATAGATAT	489
QY	1675	ttcaaaactactatgycacaaatggatagcttcccccctcatccaaagaatcaaa	1734
Db	490	TTTGGAAACTTTATATGGGAAAATATGAGATGTTTCCCACTTATATATGACCAACAGAA	549
QY	1735	aglatggagccagattatcaglygcnaatttcttglattaattlygcgcatt	1794
Db	550	GATATTTGGAGCAAAAGGATTTCTTGGAAATTTCTTAGGTGTAACCTTCGCGCTTT	609
QY	1795	atcatcatagtttttcccatggaagcagtgtttaagtgcttcatcaaaatgycataa	1854
Db	610	CTCATCATTTGTCTTTCCATATTTACTATAGTTCTGTCCATTCATCAAAAACCCGCTTGCG	669
QY	1855	gcacctgaaatacggaaatcaagttaaaaaagaatgaltcclygcnaacglttttctt	1914
Db	670	ACCACAGAAAGTAAAGCAATGTTTGGAGAGAGGGCGCTGTGCAAAATGTTCTTTT	729
QY	1915	atagcatctacgatacgatatacgcggatagccattttgtatgtaaatlttccacg	1974
Db	730	ATAGTGTCTTCGATGCCATCTGCTGGATTTCTGTATTTGTATTAATAAATCCTTTCCCT	789
QY	1975	cttcaggatgaataaccaglytaccaatacctctlygtatgatttlttctgcacat	2034
Db	790	TTCCGGGTGGAATACACAGACACAAATGACTTCTGATAGTATTTTTCCTTCCAGTT	849
QY	2035	aacagtgtcttgaacccaattctctatactctcgacacaaagccatttaagaatagatt	2094
Db	850	AACAGTGTAAATCCATCTCTATATCTCACAAACCACTTTTAAAGACAAGTTG	909
QY	2095	catcggttttgatataactacagacaaagaataatctatgacagcaaaagtccagaaca	2154
Db	910	AACAGCTGCTGCACAAACA--TCAGAGGAATCAATTTTCMAAATTAATAAAAAAAGT	966
QY	2155	tatgctccatcatcatctcgtggtga	2180
Db	967	TTATCTACATCATTTGTGTGATAGA	992

TITLE
JOURNAL
REFERENCES
AUTHORS
TITLE
JOURNAL
COMMENT

Hernandez, O., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homsi, F., Howard, B., Jia, Y., Johnson, R., Jolivet, S., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovach, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichargos, O., Lieu, C., Liu, J., Liu, W., Lounsberg, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapa, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhinney, E., McLeod, M.P., Meador, M., Mel, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokkenwo, S., Oguh, M., Okunoye, G., Ogaunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primm, E., Pu, L.L., Quiles, S., Ren, Y., Rivers, M., Rojas, A., Rojiboken, I., Rolfe, M., Ruiz, S., Saevary, G., Scherer, S., Scott, G., Shen, H., Shoohtlari, N., Stilson, I., Sodergren, E., Sonatke, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Washington, S., Williams, G., Williamson, A., Wlezyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstock, G. and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 171000)
Worley, K.C.

Direct Submission
Submitted (26-OCT-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Dec 20, 2001 this sequence version replaced gi:17064612.

AC0098607	AC0098607	171000 bp	DNA	linear	HTG 20-DEC-2001
AC0098607/c	AC0098607	171000 bp	DNA	linear	HTG 20-DEC-2001
LOCUS	Rattus norvegicus clone CH30-148B15, ***	171000 bp	DNA	linear	HTG 20-DEC-2001
DEFINITION	Rattus norvegicus clone CH30-148B15, ***	171000 bp	DNA	linear	HTG 20-DEC-2001
ACCESSION	AC0098607	171000 bp	DNA	linear	HTG 20-DEC-2001
VERSION	AC0098607.3	171000 bp	DNA	linear	HTG 20-DEC-2001
KEYWORDS	HTG; HTGS; PHASE1.	171000 bp	DNA	linear	HTG 20-DEC-2001
SOURCE	Norway rat.	171000 bp	DNA	linear	HTG 20-DEC-2001

REFERENCE
AUTHORS

1 (bases 1 to 171000)

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amarantunge,H.C., Are,J.R., Banks,T., Barbara,J., Benton,J., Blmage,K., Blankenburg,K., Bonnah,D., Bouck,J., Bowe,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Butkelt,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Day-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Dem,A.L., Ding,T., Dim,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferriguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,

```

* Consensus quality: 139190 bases at least Q40
* Consensus quality: 146928 bases at least Q20
* Consensus quality: 152581 bases at least Q20
* Estimated insert size: 137690; sum-of-contigs estimation
* Quality coverage: 0x in Q20 bases; agatose-fp estimation
* Quality coverage: 2.1x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html)
* NOTE: This is a 'working draft' sequence. It currently
* consists of 66 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 7670: contig of 7670 bp in length
* 7671 7770: gap of unknown length
* 7771 12545: contig of 4775 bp in length
* 12546 12645: gap of unknown length
* 12646 16149: contig of 3504 bp in length
* 16150 16249: gap of unknown length
* 16250 22100: contig of 5851 bp in length
* 22101 22200: gap of unknown length
* 22201 26585: contig of 4385 bp in length
* 26586 26685: gap of unknown length
* 26686 32167: contig of 5482 bp in length
* 32168 32267: gap of unknown length

```

32268	37617:	contig of 5350	bp in length
37618	37717:	gap of unknown	length
37718	43062:	contig of 5345	bp in length
43063	43162:	gap of unknown	length
43163	46609:	contig of 3447	bp in length
46610	46709:	gap of unknown	length
46710	50626:	contig of 3917	bp in length
50627	50726:	gap of unknown	length
50727	54606:	contig of 3880	bp in length
54607	54706:	gap of unknown	length
54707	57421:	contig of 2715	bp in length
57422	57521:	gap of unknown	length
57522	60497:	contig of 2976	bp in length
60498	60597:	gap of unknown	length
60598	64460:	contig of 3863	bp in length
64461	66104:	gap of unknown	length
64561	68105:	contig of 3544	bp in length
68205	71451:	gap of unknown	length
71452	71551:	gap of unknown	length
71552	74222:	contig of 2671	bp in length
74223	74322:	gap of unknown	length
74323	77237:	contig of 2915	bp in length
77238	77337:	gap of unknown	length
77338	79501:	contig of 2164	bp in length
79502	79601:	gap of unknown	length
79602	82022:	contig of 2421	bp in length
82023	82123:	gap of unknown	length
82123	84668:	contig of 2546	bp in length
84669	84768:	gap of unknown	length
84769	86880:	contig of 2112	bp in length
86881	86980:	gap of unknown	length
86981	89255:	contig of 2275	bp in length
89256	89355:	gap of unknown	length
89356	93173:	contig of 3818	bp in length
93174	93273:	gap of unknown	length
93274	96012:	contig of 2739	bp in length
96013	96112:	gap of unknown	length
96113	96812:	contig of 2500	bp in length
96813	98412:	gap of unknown	length
98413	101547:	contig of 2835	bp in length
101548	101647:	gap of unknown	length
101648	103910:	contig of 2263	bp in length
103911	104010:	gap of unknown	length
104011	107320:	contig of 3310	bp in length
107321	107420:	gap of unknown	length
107421	109694:	contig of 2274	bp in length
109695	109794:	gap of unknown	length
109795	111726:	contig of 1932	bp in length
111727	111826:	gap of unknown	length
111827	113814:	contig of 1988	bp in length
113815	113914:	gap of unknown	length
113915	115916:	contig of 1682	bp in length
115917	115966:	gap of unknown	length
115657	117489:	contig of 1793	bp in length
117490	117589:	gap of unknown	length
117590	119105:	contig of 1516	bp in length
119106	119205:	gap of unknown	length
119206	120898:	contig of 1693	bp in length
120899	120998:	gap of unknown	length
120999	123727:	contig of 2729	bp in length
123728	123827:	gap of unknown	length
123828	126119:	contig of 2292	bp in length
126120	126219:	gap of unknown	length
126220	127458:	contig of 1239	bp in length
127459	127558:	gap of unknown	length
127559	129450:	contig of 1892	bp in length
129451	129550:	gap of unknown	length
129551	133235:	contig of 2685	bp in length
132326	133235:	gap of unknown	length
132336	134307:	contig of 1972	bp in length
134308	135407:	gap of unknown	length
135409	135426:	contig of 1019	bp in length

Query Match	Best Local Similarity	6.9% 71.3% Matches	Score Pred. No. 1,4e-40; Conservative	DB 2; 0; Mismatches	Length 171000; 152; Indels	Gaps 21; Gaps
QY 1935 atgcgtgatcccaattcttgtagtgaatctcttcacgcgtcaggtagaataccag	135526	*	gap of unknown length			
DB 133291 AGGCGAGATATGCCCTTACAAATGCGACACAGCTCAATATGTTTCATCTTCTCCAGA	135527	*	contig of 226 bp in length			
QY 1995 taccaatacccttcgggtgagtgatcttattctctgcacattaaagtcgttcttgaaccaat	137792	*	contig of 191 bp in length			
DB 133231 TACTTATTTACCTTGGGGGAGATTTTATTCCTCCATCAACAGTGCCTTTGAAGCCCAAT	137793	*	gap of unknown length			
QY 2055 tctctatacctgcagcaagaacatttaagaatgatcatcgtgcgttcgtgtaacta	137892	*	contig of 2232 bp in length			
DB 133171 TATCTACACGTTTGACCACTTAACCTTTCAAGAAATGATTCATCAACTCGTGTAACCTA	137893	*	contig of 1105 bp in length			
QY 2115 cagacaaagaataatctatgacagcaaaag--tcagaaaaacataatgctccatcatcat	140124	*	gap of unknown length			
DB 133111 CAGACAAAGGAGGCTGTTTGACAGAGAAAGGACGTCAAGAAACATACATCATCATTCAT	140125	*	gap of unknown length			
QY 2172 ctgggtggaatgtyggccactgcagagagatgccaactcgagttaatgaagccgaccltt	140225	*	contig of 1713 bp in length			
DB 133051 CTGGGCTGGAATATGTCGCCCTTGCGAGGATGTTCCACAGATGCTTCAGAGCGGGATGCTTT	141938	*	gap of unknown length			
QY 2232 cacatacccccgtgaagaatgtaacgatttcacatacgaagagatcaattccattcatg	142037	*	contig of 1685 bp in length			
DB 132991 CACAGACCCCTGTGATCTGTGTGCTAATGTTTCTCGGTCACTTAAGACTCAATTCCTAATTCATA	143722	*	gap of unknown length			
QY 2292 actgactctgaatatcatcttcctgcagagataactgtygggtgcttcataggaagatt	143723	*	contig of 1293 bp in length			
DB 132931 ACTGGCTTTGGAAACCACTTATTTCTGCA-AGAAATACTGT-GAAGTGCCTTAAAGGGGTTT	145115	*	contig of 1553 bp in length			
QY 2352 actggtatgaataatgataccacaanaatlaattatataataatagctaaagataatctt	146765	*	gap of unknown length			
DB 132873 GCTGTATATG-AAGCAAAATACCACAAAATATATCTAATGAAGCAAGTTAGCTTATATGTGGCA	146766	*	contig of 1917 bp in length			
QY 2412 acaaggaacatgaggaataataaataatgactaataatgctcttcaaaaggaagaaatlaatac	148686	*	contig of 1105 bp in length			
DB 132814 CGAGG-----ACAGAAAGGTAGCTGTGCTTCATCAATCAAGAAAGCAAGCAATCACTACT	148689	*	contig of 1432 bp in length			
QY 2472 aataatgt-----ataataatagtagaactattgcatgaagaatgaagagaatcact	150311	*	contig of 1112 bp in length			
DB 132764 GACAAATGTGTCATGATATATATAGTACATA-TTTCATACGACAGATGAAGAAATCTAATTTT	150411	*	gap of unknown length			
QY 2528 cag 2530	156287	*	gap of unknown length			
	156287	*	contig of 1235 bp in length			
	157621	*	gap of unknown length			


```

Db      2112 ACCTTAACAGCTTCACTCAATGGAACAAAGATAGACTATTGGAAGAGACACATTC 2171
      893 caccctccagaagaactgagatgattaggaagtaaaagattgaaaccttcac 952
      2172 CACTTATGATACCTTAACAGTATTAGACCTGTAGACCAAGAGGTGACCGCATGTGTATA 2231
      953 ccgttatatcaagaagcgtgaagaagcgtcacaaatgaaatccttcctataatccatcc 1012
      2232 AAAACATGTTCAAAAGAGCTGAACAAATACACATTTTAAACATATCCAGAAACACATCA 2291
      1013 agaaaatcaagcaaaccaattgattatctgtcaaaactcaagtcctcagcctagaag 1072
      2292 ATTCCATAGCAAAATGAGAGCAATTAATATTTTACCCCAATGACCTTGATGATTTGCTG 2351
      1073 ggaattaaattcaaatccaaacaaagagatgtagaacctctatagaactctctcaca 1132
      2352 GCATGTGATAAAGACATTTGGTCAAAAGGTTCATGGCTTACCAAGGCTTGTTGAGC 2411
      1133 tataattagaagaattccagtaactgtggtatgacacacagtctgcagctgtaaaccaa 1192
      2412 TAAAACTGACACCTACCGGTTTGTGCTGCTCCAGAGGGTGTCAAGTGCTCAACCCA 2471
      1193 accctgatggaattcatctctagagaatctcttgcaagcatatcagaaggtatttg 1252
      2472 AACCAATGAGTTTCCCTGCTGTAAGACCTCATGTCCAAACCATGTGTGAGGGTCAGCA 2531
      1253 tctgggtgatctgcagtaactgccttgtagaacatttttgcatlttgatcgacctt 1312
      2532 TCTGGGTGCTCGGTGTCATGCGACGTGCGGAACCTTTGTCGTCATCTTTGAGAGGCTCA 2591
      1313 atatcagcttgagaacaaagcgtatgacatgtaacatctctctcgtctgctgcagact 1372
      2592 GAGATTTTCAGAGAGAGAAAGGTTCACTTTTAAATACCAACTTGGCCATGTGAGAGACT 2651
      1373 gcttaatggagaatattattatctcgatcggaagccttggaaccctgaagttcgtggagaat 1432
      2652 TTCTAATGGGGGCTACCTGTTATATCATGCCACCCGCCACCTACTACAGGGGAGTGT 2711
      1433 acataagcgtgcagcgtgtagatgagagactacatctgcaactgtagaactcttg 1492
      2712 ACATTCAGCCACGATGAGAACTGGAACAGAGTGGACTCTGTGCTGCTGAGTTTGGCTCT 2771
      1493 ccaatctgtccacagaagatcagtttaactctgtaacattctgacatctggaataaataca 1552
      2772 CCACATTTTCTAGTAATATATCATGCTTTCCTGTCACCATCATACACTTGACCGGCTCA 2831
      1553 tctgcatctgtctatcctctttagatgtgtgagacctggaataatgcagaaatacagttc 1612
      2832 TCTGCATCTCTGTTTCCCTGAGAGAACTGCACTTGCCCTGGAACAGGCCATCATTTGTCA 2891
      1613 tgaattcatcttgatctgactggtttatagtgcttcaatctgagcattgagaataaagaat 1672
      2892 TGAAGCTGATATAGGTGCTCGGTCTCTGCTGCTGCTGCTTACCTCTGTTTAGATTTTTCAT 2951
      1673 ttctcaaaaactactatggaacaaatgagatgcttccctctctcaatcagaagaatacag 1732
      2952 ACTTGAAGAACTCTACGCGCGCTGTGAGTATGCTGCGCCCTGCA---GTAATACCTCA 3008
      1733 aaagtgtggagcccaagattatcagtgagcaattttctctgtatataattggccgcgat 1792
      3009 ACAGGGGGCCAGGCTGGAGTATCTGTTGTTGTTCTTCATCCCTTCATACCTGCTGTGAT 3068
      1793 ttaatcatagatctttctcctatgagaagcatgtttatatagtgtaacaaagtcataa 1852
      3069 TCGTGTCTCATAGCGCTCTCTTACCTTTGATGTTCATAGTACGCAAAAGACCCGAGAGTG 3128
      1853 cagcaactgaataacagaaatcaagttaaaaaagagaatgactccttgccaaagcttttctc 1912
      3129 CA-----GTCAAGGACTGCTGAGAGCAAAATGACAAATGCGCAGAGGATGACT 3182
      1913 ttatagttattactgactgactatgctgtagataaccaattttgtagtgaattcttctc 1972

```

```

Db      3183 TGATTGTCATGACTGACTTTTCTGCTGGTCCCATCATTTGTCTTGGCTTTGTTTC 3242
      1973 tgcctcaggtagaataaccaggtaccataaccctctgggtagatgtattatctgcaca 2032
      3243 TACCGGCTGCTAGGCGGATGATCAGTGTATGCAATGATGATGCTGTGTTGCTCTCCCG 3302
      2033 ttacagatgcttgaaccaatctctatactctgacacaaagaccat 2081
      3303 TGAACCTGCGACCAACCCGTCATCTACACCTCTCCACAGCGCATTT 3351

RESULT 11
AL138708 170522 bp DNA linear PRI 01-FEB-2001
LOCUS Human DNA sequence from clone RP11-432E15 on chromosome 13 Contains
DEFINITION STSS, GSSs and a Cpg island, complete sequence.
ACCESSION AL138708.17 GI:11137618
VERSION AL138708.17 GI:11137618
KEYWORDS HTG; Cpg Island.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 170522)
REFERENCE 1 (bases 1 to 170522)
AUTHORS Philimore,B.
TITLE Direct Submission
JOURNAL Submitted (01-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
COMMENT On Nov 10, 2000 this sequence version replaced gi:11121335.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
This sequence was generated from part of bacterial clone contigs of human
chromosome 13, constructed by the Sanger Centre Chromosome 13
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr13
This sequence is the entire insert of clone RP11-432E15. The true
left end of clone RP11-154I23 is at 137612 in this sequence. The
true right end of clone RP11-95N14 is at 13246 in this sequence.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated repeat sequence elements. Where the sequence is
ambiguous, there is an annotation using the 'unsure' feature key.
RP11-432E15 is from the library RPCR-11.2 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBAC3.6.
FEATURES
source 1..170522
location/Ouallifiers
1..170522
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="13"
/clone="RP11-432E15"
/clone.lib="RPCR-11.2"
19..289
/misc-feature
/note="match: GSS: Em:A0549338"
194..299
/repeat_region
/note="53 copies 2 mer tg 71% conserved"
208..299
/repeat_region
/note="23 copies 4 mer tgtg 77% conserved"
1456..1750
/repeat_region
/note="AlusX repeat: matches 3..300 of consensus"
2078..2183

```

```
misc_feature      /note="MIR repeat: matches 6. .114 of consensus"
                  2600. .3122
/note="match: GSS: Em:AQ529216"
repeat_region     /note="AluSg repeat: matches 1. .302 of consensus"
                  2646. .2940
/note="AluSg repeat: matches 1. .296 of consensus"
                  3590. .3894
/note="AluSg repeat: matches 1. .296 of consensus"
                  3926. .4462
/note="L2 repeat: matches 2233. .2750 of consensus"
                  4658. .4805
/note="L2 repeat: matches 2355. .2503 of consensus"
                  4911. .5224
/note="AluSg repeat: matches 1. .312 of consensus"
                  5237. .5332
/note="L2 repeat: matches 2575. .2694 of consensus"
                  5788. .6085
/note="AluJ repeat: matches 3. .301 of consensus"
                  7355. .7747
/note="sequence from overlapping clone bA15909
                  (A1136106).Assembly confirmed by restriction digest."
repeat_region     /note="L2 repeat: matches 2 mer at 63% conserved"
                  7388. .7487
/note="50 copies 2 mer at 63% conserved"
                  7396. .7487
/note="23 copies 4 mer at 66% conserved"
                  7884. .7915
/note="MER20 repeat: matches 30. .62 of consensus"
                  7916. .8219
/note="AluJb repeat: matches 1. .302 of consensus"
                  8220. .8376
/note="MER20 repeat: matches 62. .210 of consensus"
                  8975. .8985
/note="sequence from overlapping clone
                  bA15909(A1136106).Assembly confirmed by restriction
                  digest."
repeat_region     /note="11 copies 4 mer ca 100% conserved"
                  9078. .9121
/note="complement(9603. .10104)"
                  /note="match: GSS: Em:AQ569313"
                  11234. .11331
/note="sequence from overlapping clone bA15909
                  (A1136106).Assembly confirmed by restriction digest."
misc_feature      /note="match: GSS: Em:AQ486928"
                  12477. .12601
/note="MIR repeat: matches 15. .138 of consensus"
                  12713. .12980
/note="sequence from overlapping clone bA15909
                  (A1136106).Assembly confirmed by restriction digest."
repeat_region     /note="MIR repeat: matches 38. .188 of consensus"
                  13372. .13417
/note="complement(13670. .13919)"
                  /note="match: GSS: Em:AQ021664"
                  /note="complement(13740. .14130)"
                  /note="match: GSS: Em:AQ032630"
                  13771. .14227
/note="sequence from overlapping clone bA15909
                  (A1136106).Assembly confirmed by restriction digest."
repeat_region     /note="L2 repeat: matches 2515. .2750 of consensus"
                  13864. .14087
/note="L2 repeat: matches 1. .244 of consensus"
                  16061. .16342
/note="MER1B repeat: matches 1. .244 of consensus"
                  16343. .16655
/note="AluSx repeat: matches 1. .312 of consensus"
                  16656. .16680
/note="MER1B repeat: matches 244. .305 of consensus"
                  /complement(17019. .17275)
                  /note="match: GSS: Em:AQ452676"
                  17070. .17578
/note="match: GSS: Em:AQ511484"
                  17087. .17515
/note="match: GSS: Em:AQ0817416"
                  17278. .17794
/note="match: GSS: Em:AQ358157"

repeat_region     18301. .18381
/misc_feature     /note="L1MB2 repeat: matches 6082. .6166 of consensus"
                  complement(18563. .19010)
                  /note="match: GSS: Em:AQ242897"
                  20403. .20559
/note="MER20 repeat: matches 1. .153 of consensus"
                  20870. .21069
/note="100 copies 2 mer tt 69% conserved"
                  20937. .21076
/note="35 copies 4 mer ttcc 72% conserved"
                  21080. .21392
/note="AluJb repeat: matches 1. .310 of consensus"
                  21395. .21538
/note="MIR repeat: matches 94. .245 of consensus"
                  23716. .23960
/note="L2 repeat: matches 2381. .2617 of consensus"
                  23961. .24213
/note="AluSx repeat: matches 47. .299 of consensus"
                  24214. .24342
/note="L2 repeat: matches 2617. .2747 of consensus"
                  25636. .25785
/note="MER5A repeat: matches 32. .189 of consensus"
                  25638. .25810
/note="MER5A repeat: matches 17. .189 of consensus"
                  26013. .26227
/note="MIR repeat: matches 10. .256 of consensus"
                  28426. .28484
/note="L2 repeat: matches 2641. .2700 of consensus"
                  28639. .28960
/note="AluSg repeat: matches 1. .310 of consensus"
                  29689. .29861
/note="L2 repeat: matches 2006. .2166 of consensus"
                  complement(29720. .30187)
                  /note="match: GSS: Em:AQ317477"
                  29901. .30075
/note="MER5A repeat: matches 1. .184 of consensus"
                  /note="complement(30119. .30414)"
                  /note="match: GSS: Em:AQ230143"
                  32462. .32774
/note="AluSx repeat: matches 1. .312 of consensus"
                  32807. .32877
/note="L2 repeat: matches 2635. .2707 of consensus"
                  35365. .35806
/note="L2 repeat: matches 2248. .2742 of consensus"
                  complement(35852. .36413)
                  /note="match: GSS: Em:AQ284653"
                  match: STS: Em:G53501"
                  complement(36043. .36433)
                  /note="match: GSS: Em:AQ816559"
                  36420. .36469
/note="MIR repeat: matches 96. .148 of consensus"
                  36429. .36961
/note="match: GSS: Em:AQ394779"
                  36452. .36845
/note="match: GSS: Em:AQ212353"
                  39203. .39242
/note="20 copies 2 mer ca 100% conserved"
                  39250. .39648
/note="L2 repeat: matches 2307. .2710 of consensus"
                  39703. .39750
/note="24 copies 2 mer ca 75% conserved"
                  39859. .39976
/note="L1MC3 repeat: matches 7621. .7739 of consensus"
                  39988. .40283
/note="AluSx repeat: matches 1. .296 of consensus"
                  40873. .41144
/note="AluY repeat: matches 39. .310 of consensus"

Query Match      5.2%: Score 187.4; DB 9; Length 170522;
Best Local Similarity 64.8%: Pred. NO. 2.9e-28;
Matches 278; Conservative 0; Mismatches 151; Indels 0; Gaps 0;
```

QY 1349 tcaattctctgtgtccgactgcttaatgggaatatattatctgtatcgagagct 1408


```
misc_feature      65270..75508
                  /note="assembly_fragment:00461
                  fragment_chain:2"
misc_feature      75609..80749
                  /note="assembly_fragment:01454
                  fragment_chain:2"
misc_feature      80850..101473
                  /note="assembly_fragment:00059
                  fragment_chain:2"
misc_feature      101574..108695
                  /note="assembly_fragment:01172
                  fragment_chain:3"
misc_feature      108796..120411
                  /note="assembly_fragment:01458
                  fragment_chain:3"
misc_feature      120512..125324
                  /note="assembly_fragment:00440"
misc_feature      125425..129330
                  /note="assembly_fragment:00447"
misc_feature      129431..135864
                  /note="assembly_fragment:00564"
misc_feature      135965..156228
                  /note="assembly_fragment:00960"
misc_feature      156329..161861
                  /note="assembly_fragment:01066.0"
misc_feature      161962..164531
                  /note="assembly_fragment:01287"
misc_feature      164632..169371
                  /note="assembly_fragment:01319"
misc_feature      169472..176565
                  /note="assembly_fragment:01453"
BASE COUNT      52990 a 35385 c 33749 g 52527 t 1914 others
ORIGIN
```

```
Query Match      5.2%; Score 187.4; DB 2; Length 176565;
Best Local Similarity 64.8%; Pred. No. 2.9e-28;
Matches 278; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 1349 tcaattctctctgtgtccgacgtcgttaagggaataattatctgtatcgaggagct 1408
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 97765 TCAATTCCTCCACAGAGTCTATTCCTGATGCGGCTTACTGTTCTTTGGCAATTT 97824

QY 1409 ttgacccaagtttcgtgtgagaataacaatgaacgtgcagctgtgtgtgagagtaactc 1468
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 97825 TCGATATATAATACCGAGGCGAGTATCGAAGATATGCTTCTGCTGATGGAGCGTGC 97864

QY 1469 attgtcagctgtgagatccttggccattctgtccacagaagatcacgtttactgttaa 1528
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 97885 AGTGGCGGCTCATGGGGTTCCTGGCCATGCGTCCACGGAAGTCTGTTCTGCTACTGCA 97944

QY 1529 catttcagcatctggaanaataacatcgtcatgtctatccctttagatgtgtgagacctg 1588
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 97945 CCTACTTCACTTGTGGAGAGTCTCCGTGTCATGTTCTCCCTTCAGTAACATTCGACCTG 98004

QY 1589 gaaaatgagaacaattacaattcgtatcattcattcattcattcattcattcattcattc 1648
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 98005 GAAAGCGGACAGCTCACTCACTCACTTGCATTCGATGCGGAGGATTTTAAATAGCG 98064

QY 1649 tcaattcattgagcaataagaattttcaaaactactactatgcccacaaatgagatgtct 1708
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 98065 TAATTCATTTTGGAAATAGAGTATTTTGGAACTTTTATGCGGAAATATGAGATATGTT 98124

QY 1709 tccctctcattcagaagatcagaagaatlatgagccacagattatcactcagtggaattt 1768
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 98125 TCCCACTTATATGACCAACAGAGATATTTGGAAGCAAGGATATCTCTTGAATTT 98184

QY 1769 ttcttggtta 1777
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 98185 TCCTAGGTA 98193
```

RESULT 13

```
AC098607
LOCUS              171000 bp      DNA      linear      HTG 20-DEC-2001
DEFINITION        Rattus norvegicus clone CH230-148B15, *** SEQUENCING IN PROGRESS
ACCESSION         AC098607
VERSION           AC098607.3 GI:17973780
KEYWORDS          HTG; HTGS PHASE1.
SOURCE            Norway rat.
ORGANISM          Rattus norvegicus
                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
                  Rattus.
REFERENCE
AUTHORS           Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
                  Albrooks,S.L., Amaratunga,H.C., Are,J.R., Banks,T., Barbata,J.,
                  Benton,J., Bimaje,K., Blankenburg,K., Bonnin,D., Bouck,D.,
                  Bowie,S., Briteva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
                  Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Caron,T.F.,
                  Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
                  Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
                  Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
                  Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
                  Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
                  Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
                  Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
                  Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
                  Garza,N., Gill,R., Gorrell,J.H., Guerrero,M., Gunaratne,P., Hale,S.,
                  Hamilton,K., Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A.,
                  Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
                  Hollins,B., Homai,F., Howard,S., Huber,J., Huiyk,S., Hume,J.,
                  Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
                  Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
                  Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
                  Lewis,L., Li,J., Li,Z., Lichtenarge,O., Lieu,C., Liu,J., Liu,W.,
                  Louisaged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
                  Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
                  Martinez,E., Massey,E., Mawhinney,E., Mcleod,M.P., Meador,M.,
                  Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
                  Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
                  Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokkenko,S.,
                  Oguh,M., Okwuonu,G., Oregunye,N., Oyedero,R., Pace,A., Payton,B.,
                  Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
                  Ruiz,S., Ren,Y., Rivers,M., Rojas,A., Rojibokan,I., Rolfe,M.,
                  Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,
                  Sisson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H.,
                  Stone,H., Sutton,A., Svatek,A., Tabot,P., Tamerisa,A., Tamerisa,K.,
                  Tang,H., Tansey,J., Taylor,C., Taylor,T., Tellod,B., Thomas,N.,
                  Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalona,D., Vinson,R.,
                  Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
                  Watlington,S., Williams,G., Williamson,A., Wleczky,R., Wooden,S.,
                  Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
                  Weinstein,G. and Gibbs,R.
TITLE             Unpublished
JOURNAL           Direct Submission
AUTHORS           2 (bases 1 to 171000)
REFERENCE         Worley,K.C.
TITLE             Direct Submission
JOURNAL           Submitted (26-OCT-2001) Human Genome Sequencing Center, Department
                  of Molecular and Human Genetics, Baylor College of Medicine, One
                  Baylor Plaza, Houston, TX 77030, USA
                  On Dec 20, 2001 this sequence version replaced gi:17064612.
COMMENT           ----- Genome Center
                  Center: Baylor College of Medicine
                  Center code: BCM
                  Web site: http://www.hgsc.bcm.tmc.edu/
                  Contact: hgsc-help@bcm.tmc.edu
                  ----- Project Information
                  Center project name: GTCC
                  Center clone name: CH230-148B15
                  ----- Summary Statistics
                  Assembly program: Phrap; version 0.990329first call to
                  findphraplist
                  Consensus quality: 139190 bases at least Q40
```


Contact: saplens@watson.wustl.edu
 ----- Summary Statistics
 Center project name: H_NH0719M18

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
 The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frengen, E., Tatenno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBAC3.6
 NEIGHBORING SEQUENCE INFORMATION:
 Actual start of this clone is at base position 1 of RP11-719M18; actual end is at base position 190122 of RP11-719M18.

FEATURES

source

1. 190122
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="4"
 /map="4"
 /clone="RP11-719M18"
 /clone_lib="RPCI-11"
 505. 553
 repeat_region
 /rpt_family="AT_rich"
 643. 975
 repeat_region
 /rpt_family="MaLR"
 972. 1001
 repeat_region
 /rpt_family="(TG)n"
 1015. 1216
 repeat_region
 /rpt_family="Alu"
 1685. 1853
 repeat_region
 /rpt_family="MIR"
 2241. 2304
 repeat_region
 /rpt_family="AT_rich"
 2429. 2833
 repeat_region
 /rpt_family="LI"
 3404. 4027
 repeat_region
 /rpt_family="ERV1"
 4032. 4422
 repeat_region
 /rpt_family="ACHobo"
 5753. 5842
 repeat_region
 /rpt_family="LI"
 5843. 6217
 repeat_region
 /rpt_family="MaLR"
 6225. 6355
 repeat_region
 /rpt_family="MaLR"
 6353. 6733
 repeat_region
 /rpt_family="LI"
 6734. 7212
 repeat_region
 /rpt_family="MER1_type"

repeat_region 7213. 7305
 /rpt_family="LI"
 repeat_region 7809. 8000
 /rpt_family="MIR"
 repeat_region 8274. 8748
 /rpt_family="MaLR"
 repeat_region 8955. 9081
 /rpt_family="L2"
 repeat_region 9289. 9579
 /rpt_family="Alu"
 repeat_region 10261. 10360
 /rpt_family="MaLR"
 repeat_region 10446. 10611
 /rpt_family="ERV1"
 repeat_region 10614. 11158
 /rpt_family="ERV1"
 repeat_region 11280. 11389
 /rpt_family="MER1_type"
 repeat_region 11400. 11819
 /rpt_family="ERV1"
 repeat_region 12071. 12388
 /rpt_family="Alu"
 repeat_region 12359. 12405
 /rpt_family="(A)n"
 repeat_region 12620. 12973
 /rpt_family="MaLR"
 repeat_region 13201. 13221
 /rpt_family="AT_rich"
 repeat_region 13207. 13368
 /rpt_family="Alu"
 repeat_region 13863. 14381
 /rpt_family="ERV1"
 repeat_region 15625. 15659
 /rpt_family="AT_rich"
 repeat_region 16105. 16445
 /rpt_family="L2"
 repeat_region 16584. 16945
 /rpt_family="LI"
 repeat_region 16934. 17030
 /rpt_family="(CATATA)n"
 repeat_region 19292. 19314
 /rpt_family="(T)n"
 repeat_region 19843. 19954
 /rpt_family="T-rich"
 repeat_region 24855. 24934
 /rpt_family="LI"
 repeat_region 24959. 25593
 /rpt_family="ACHobo"
 repeat_region 25594. 25951
 /rpt_family="MaLR"
 repeat_region 25952. 26253
 /rpt_family="ACHobo"
 repeat_region 26281. 26515
 /rpt_family="LI"
 repeat_region 26508. 26530
 /rpt_family="AT_rich"
 repeat_region 26935. 26986
 /rpt_family="(TTTA)n"
 repeat_region 26985. 27435
 /rpt_family="LI"
 repeat_region 27628. 27653
 /rpt_family="AT_rich"
 repeat_region 27761. 27781
 /rpt_family="(TA)n"
 repeat_region 27781. 28057
 /rpt_family="Alu"
 repeat_region 28062. 28082
 /rpt_family="AT_rich"
 repeat_region 28297. 28358
 /rpt_family="CT-rich"
 repeat_region 28322. 28626
 /rpt_family="LI"
 repeat_region 28741. 28867

```
repeat_region /rpt_family="Alu"
28875. .28931 /rpt_family="AT_rich"
repeat_region 28930. .29909 /rpt_family="L1"
repeat_region 29915. .30295 /rpt_family="L1"
repeat_region 30351. .30585 /rpt_family="MALR"
repeat_region 30864. .30953 /rpt_family="MIR"
repeat_region 31260. .31675 /rpt_family="L1"
repeat_region 31678. .32066 /rpt_family="L1"
```

Query Match 4.8; Score 173; DB 9; Length 190122;

Best Local Similarity 97.28; Pred. No. 2.5e-25; Matches 176; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
QY 1 ctgcttgaactgtaagattgagacagaaatagcacacaaccactgtagctgatatg 60
|||
Db 124029 CTGCTTGTAACTGCTAAGATGAGACAGAAATAGCACACACCACTGAGCTGATG 124088
|||
QY 61 cgattcagaacccaagccaatttgcacttcaatcaatcagttgctcagatagaag 120
|||
Db 124089 CGATTGAGAAACCAAGACCAATTTGCTCCTTCATTAATGAGTGCAGATAGAG 124148
|||
QY 121 gaaatgacatcgtctgctctcttcaacatctaattttggaaaatatcttccat 180
|||||
Db 124149 GAAATGACATCTGCTCTCTCTCTACATCTTAATTTGGAAAATATTTGATAGTAT 124208
|||||
QY 181 g 181
|
Db 124209 G 124209
```

Search completed: September 5, 2002, 08:34:23
Job time: 20183 sec

[illegible]

|||||
Db 1801 atagtttttccatggaagatggttttatagtgltcatcaagtgcataacagcaact 1860
QY 1861 gaatacaggaatacaagttaaaaaagagaagaatcctctgcacaaagttcttcttatagta 1920
Db 1861 gaatacaggaatacaagttaaaaaagagaagaatcctctgcacaaagttcttcttatagta 1920
QY 1921 tttcctatgcatatgctgtgatagccattttgttagtaattcttctccttcag 1980
Db 1921 tttcctatgcatatgctgtgatagccattttgttagtaattcttctccttcag 1980
QY 1981 gtagaatacaccaggttacataaacctcttggttagtgaatttttattcctgcacattaacagt 2040
Db 1981 gtagaatacaccaggttacataaacctcttggttagtgaatttttattcctgcacattaacagt 2040
QY 2041 gctttgaacccaattcctctatactctgcacaaagaccatttaagaataatgattcctcg 2100
Db 2041 gctttgaacccaattcctctatactctgcacaaagaccatttaagaataatgattcctcg 2100
QY 2101 ttttggataactacagacaagaanaatctatgagacagcaaggctcagaanaacataatgct 2160
Db 2101 ttttggataactacagacaagaanaatctatgagacagcaaggctcagaanaacataatgct 2160
QY 2161 ccaatcatcactctggtggaanaatggtgccaactgcagagagatgcacacctgattatgaag 2220
Db 2161 ccaatcatcactctggtggaanaatggtgccaactgcagagagatgcacacctgattatgaag 2220
QY 2221 ccggaactcttccacataccctctgtggaatgtcactgatttccataacagagactcaat 2280
Db 2221 ccggaactcttccacataccctctgtggaatgtcactgatttccataacagagactcaat 2280
QY 2281 tccattcatgacgactcctaaattcaattcttcgcagagaatactgtgggggtgcttc 2340
Db 2281 tccattcatgacgactcctaaattcaattcttcgcagagaatactgtgggggtgcttc 2340
QY 2341 atgagggatttactggtatgaaaaatgaaatacacaaaaatattatataataatagctaa 2400
Db 2341 atgagggatttactggtatgaaaaatgaaatacacaaaaatattatataataatagctaa 2400
QY 2401 ataaatttttacaagagacatgagaanaaaaaataaaatgactaaagctcttcaaaaggaa 2460
Db 2401 ataaatttttacaagagacatgagaanaaaaaataaaatgactaaagctcttcaaaaggaa 2460
QY 2461 gtaattatcaataaagtatataatatagtagagacatttgcataagaataataagagaa 2520
Db 2461 gtaattatcaataaagtatataatatagtagagacatttgcataagaataataagagaa 2520
QY 2521 tctactcgaataacatcatcatcttcttctaacaatgcatattatgagtaaccactactat 2580
Db 2521 tctactcgaataacatcatcatcttcttctaacaatgcatattatgagtaaccactactat 2580
QY 2581 gtgcataagcatgcaataatagtcctggaagtacagtgacgagaaaccttcaactcgtlaga 2640
Db 2581 gtgcataagcatgcaataatagtcctggaagtacagtgacgagaaaccttcaactcgtlaga 2640
QY 2641 taggttttaagacaagaagactatatacaaaagtcactgcagattcctagtttaaaagttag 2700
Db 2641 taggttttaagacaagaagactatatacaaaagtcactgcagattcctagtttaaaagttag 2700
QY 2701 cttaacccgtcatgtgcatacgcagaagaatcctagacatttaataaaggtttaaaagt 2760
Db 2701 cttaacccgtcatgtgcatacgcagaagaatcctagacatttaataaaggtttaaaagt 2760
QY 2761 ttggaatactcaagtgtatgtgcatacagaagaatgctcgtactgtttgcaaaaataatctc 2820
Db 2761 ttggaatactcaagtgtatgtgcatacagaagaatgctcgtactgtttgcaaaaataatctc 2820
QY 2821 tgttttaagaatcatcttactctctttaaagttccatacaacttgagagcaacaacac 2880
Db 2821 tgttttaagaatcatcttactctctttaaagttccatacaacttgagagcaacaacac 2880
QY 2881 atattatataaagaatgcttgcatagaacccaacacagcaactcttgcagact 2940
Db 2881 atattatataaagaatgcttgcatagaacccaacacagcaactcttgcagact 2940

Db 2881 atattatataaagaatgcttgcatagaacccaacacagcaactcttgcagact 2940
QY 2941 tctgtcccaagtttctcttcttgaatgaacatcatcatatggaatttgaatagaga 3000
Db 2941 tctgtcccaagtttctcttcttgaatgaacatcatcatatggaatttgaatagaga 3000
QY 3001 gtatgagtaacgagagaagtgatgacgaanaacataagaatgataaacattacatta 3060
Db 3001 gtatgagtaacgagagaagtgatgacgaanaacataagaatgataaacattacatta 3060
QY 3061 gtggaactccttgaataaalccttgatgtgcattgactgaatttccaacaaagatgcc 3120
Db 3061 gtggaactccttgaataaalccttgatgtgcattgactgaatttccaacaaagatgcc 3120
QY 3121 aagacaaaaggcttccaacaaacgctgctgttttaagaacaaactaagtgtttaat 3180
Db 3121 aagacaaaaggcttccaacaaacgctgctgttttaagaacaaactaagtgtttaat 3180
QY 3181 tcacccactttagatgggtgatgttattgtgtgtgaataatcccaagtaagcagttaaa 3240
Db 3181 tcacccactttagatgggtgatgttattgtgtgtgaataatcccaagtaagcagttaaa 3240
QY 3241 aggaanaaagctggaatgacatgattcgaagaacttaattcaggaaggaaggctcgtta 3300
Db 3241 aggaanaaagctggaatgacatgattcgaagaacttaattcaggaaggaaggctcgtta 3300
QY 3301 tgtacacatttcaatttaagaagaanaatcttcttcaagaataatgacttaattctctt 3360
Db 3301 tgtacacatttcaatttaagaagaanaatcttcttcaagaataatgacttaattctctt 3360
QY 3361 gcactgcagacgltgagatataactctttaaactagtgtctctccttaactcctcag 3420
Db 3361 gcactgcagacgltgagatataactctttaaactagtgtctctccttaactcctcag 3420
QY 3421 tattagnaatttttgccttcaataatgtaaaacctttaaagcaggaagaagaatgtttca 3480
Db 3421 tattagnaatttttgccttcaataatgtaaaacctttaaagcaggaagaagaatgtttca 3480
QY 3481 gatagtttcaaatnccaaanaatggttgcacacaaataactcggatcnaaccataat 3540
Db 3481 gatagtttcaaatnccaaanaatggttgcacacaaataactcggatcnaaccataat 3540
QY 3541 gccctattgaaatatatagttgataagnttgcctgtaaaacc 3584
Db 3541 gccctattgaaatatatagttgataagnttgcctgtaaaacc 3584

RESULT 2
AA25345
ID AA25345 standard; cDNA; 2467 BP.
XX
AC AA25345;
XX
DT 20-DEC-1999 (first entry)
XX
DE Human LGR7 Long form nucleotide sequence.
XX
XX Human; LGR4; LGR5; LGR7; G-protein coupled receptor; gene therapy;
KW extracellular Leucine rich repeat region; mapping; identification; ss.
OS Homo sapiens.
XX
PN W09948921-A1.
XX
PD 30-SEP-1999.
XX
PE 25-MAR-1999; 99WO-US06573.
XX
PR 26-MAR-1998; 98US-0079501.
XX
PA (STRD) UNIV LEI LAND STANFORD JUNIOR.
PA (ORGA) ORGANON NV.
XX

PI Hsueh AJW, Hsu SY, Liang S, Van Der Spek PJ;
XX WPI: 1999-591074/50.
DR P-PSDB: AAY42170.
XX New G-protein coupled receptors, useful for identifying their own
PI ligands -
XX
PS Claim 4; Fig 3; 54pp; English.
XX The present sequence encodes the human G-protein coupled receptor
CC having extracellular leucine rich repeat regions, designated IGR7 long
CC form. The IGR4, IGR5 and IGR7 proteins are used to identify ligands for
CC the receptor. The polypeptides and/or polynucleotides are also useful
CC for homologous or related genes, producing compositions that modulate
CC the expression or function of the receptors, gene therapy, mapping
CC functional regions of the receptors, studying associated physiological
CC pathways, in vivo prophylactic and therapeutic purposes, as immunogens
CC for producing antibodies, and for identifying biologically active
CC agents. The polypeptides contain a G-protein coupled seven
CC transmembrane region and a leucine rich repeat extracellular domain.
CC These regions capture and facilitate optimal orientation of its ligand.
CC The proteins are also expressed in diverse tissues.
XX
SQ Sequence 2467 BP; 747 A; 487 C; 474 G; 759 T; 0 other;

Query Match 59.1%; Score 2119.4; DB 20; Length 2467;
Best Local Similarity 95.2%; Pred. No. 0;
Matches 2281; Conservative 0; Mismatches 11; Indels 105; Gaps 5;

QY 1 cgcgttgaactgtctaagattgacagacagaataagacacaacacactgtgactatg 60
DB |||||||ttgaactgtctaagattgacagacagaataagacacaacacactgtgactatg 130
71 cgcgttgaactgtctaagattgacagacagaataagacacaacacactgtgactatg 130
QY 61 cgattcagaacaacagaacaaatttgcctcacttcaattcaatgaattgctcagaatgaag 120
DB |||||||ttgaactgtctaagattgacagacagaataagacacaacacactgtgactatg 190
131 cgattcagaacaacagaacaaatttgcctcacttcaattcaatgaattgctcagaatgaag 190
QY 121 gaatatcacatcgttctgtctctctctcactcacttaattttgnaaataattttctcat 180
DB |||||||ttgaactgtctaagattgacagacagaataagacacaacacactgtgactatg 250
191 gaatatcacatcgttctgtctctctctcactcacttaattttgnaaataattttctcat 250
QY 181 gggggtgagacagagatgcaagtgctccctgtgctattccccctgtgggaacacacaag 240
DB |||||||ttgaactgtctaagattgacagacagaataagacacaacacactgtgactatg 310
251 gggggtgagacagagatgcaagtgctccctgtgctattccccctgtgggaacacacaag 310
QY 241 tgccttcacagctcctcctcactgtaacggtgtgacgactgcgggaatcagcgatgag 300
DB |||||||ttgaactgtctaagattgacagacagaataagacacaacacactgtgactatg 370
311 tgccttcacagctcctcctcactgtaacggtgtgacgactgcgggaatcagcgatgag 370
QY 301 gacaactgtg----- 310
DB |||||||ttgaactgtctaagattgacagacagaataagacacaacacactgtgactatg 430
371 gacaactgtgagacaacaatgtagtgctcatgcaatttgacaataatttgcacgttac 430
QY 311 -----tggtggtttt 320
DB |||||||ttgaactgtctaagattgacagacagaataagacacaacacactgtgactatg 490
431 tacaacatgacttcccaatatccttctgagcgagaacacactgtgactatgctgcgttct 490
QY 321 gtgcacagtcg-attgtcttgcc--aggtctgagactgactg--gtatgaacca--tttagc 375
DB |||||||ttgaactgtctaagattgacagacagaataagacacaacacactgtgactatg 550
491 gtgcacagtcgaaatgctcttcccaagtcctgagactgtgacgttgatgnaacaaatttagc 550
QY 376 agtcttcacagcttcttccaatgtgactgcaatgtcacttcagtggaacttaataaga 435
DB |||||||ttgaactgtctaagattgacagacagaataagacacaacacactgtgactatg 610
551 gctgttcacagcttcttccaatgtgactgcaatgtcacttcagtggaacttaataaga 610
QY 436 aagcttcctcctgactgtcttccaagaatataatgattcttccaagaactgcggactcaaac 495
DB |||||||ttgaactgtctaagattgacagacagaataagacacaacacactgtgactatg 670
611 aagcttcctcctgactgtcttccaagaatataatgattcttccaagaactgcggactcaaac 670
QY 496 aataagattacatcatcatcatatgattgcttcagagacggaataagacttaccactg 555
DB |||||||ttgaactgtctaagattgacagacagaataagacacaacacactgtgactatg

DB 671 aataagattacatcatcatcatatgattgcttcagagacggaataagacttaccactg 730
QY 556 tatctcagtcatacagaataacacttctctgaaagcgggtgttttggaaatcttccaaga 615
DB |||||||ttgaactgtctaagattgacagacagaataagacacaacacactgtgactatg 790
731 tatctcagtcatacagaataacacttctctgaaagcgggtgttttggaaatcttccaaga 790
QY 616 ctgaatgtgcataatgtgaagataatcacctcagctgcaatttccccaacatttat 675
DB |||||||ttgaactgtctaagattgacagacagaataagacacaacacactgtgactatg 850
791 ctgaatgtgcataatgtgaagataatcacctcagctgcaatttccccaacatttat 850
QY 676 ggactaaattccttatctctctagctcgtatgataaacytccctacccgttttactgact 735
DB |||||||ttgaactgtctaagattgacagacagaataagacacaacacactgtgactatg 910
851 ggactaaattccttatctctctagctcgtatgataaacytccctacccgttttactgact 910
QY 736 aaactctctgtcaacacatgccaagacttaacttggcggacactggaagcaacacatac 795
DB |||||||ttgaactgtctaagattgacagacagaataagacacaacacactgtgactatg 970
911 aaactctctgtcaacacatgccaagacttaacttggcggacactggaagcaacacatac 970
QY 796 cataatttaagaattgtacttattctctgcaagtaatttaactggttttagtgatgag 855
DB |||||||ttgaactgtctaagattgacagacagaataagacacaacacactgtgactatg 1030
971 cataatttaagaattgtacttattctctgcaagtaatttaactggttttagtgatgag 1030
QY 856 aaaaaaataatcaatcaacttaaatgaaaaatcttgcacacttccagaactgtatga 915
DB |||||||ttgaactgtctaagattgacagacagaataagacacaacacactgtgactatg 1090
1031 aaaaaaataatcaatcaacttaaatgaaaaatcttgcacacttccagaactgtatga 1090
QY 916 ttgatttaggaagtaataagaattgaaaaattctccacgcgttatattcaagactgaa 975
DB |||||||ttgaactgtctaagattgacagacagaataagacacaacacactgtgactatg 1150
1091 ttgatttaggaagtaataagaattgaaaaattctccacgcgttatattcaagactgaa 1150
QY 976 gagctgtcaaatgtaatcttctccataatccaatccagaanaattccaagcaacaatt 1035
DB |||||||ttgaactgtctaagattgacagacagaataagacacaacacactgtgactatg 1210
1151 gagctgtcaaatgtaatcttctccataatccaatccagaanaattccaagcaacaatt 1210
QY 1036 gattatcttgcaaacactcaagctctcagcctcagaaggaattgaaattccaaitccaa 1095
DB |||||||ttgaactgtctaagattgacagacagaataagacacaacacactgtgactatg 1270
1211 gattatcttgcaaacactcaagctctcagcctcagaaggaattgaaattccaaitccaa 1270
QY 1096 caaagatgtttgaagacttcttgaactcctcctcacaatatatttgaagaattccaag 1155
DB |||||||ttgaactgtctaagattgacagacagaataagacacaacacactgtgactatg 1330
1271 caaagatgtttgaagacttcttgaactcctcctcacaatatatttgaagaattccaag 1330
QY 1156 tgtgggtatgcaacacatgctcgagcgtgtaaaccaacaacactgataattcattctca 1215
DB |||||||ttgaactgtctaagattgacagacagaataagacacaacacactgtgactatg 1390
1331 tgtgggtatgcaacacatgctcgagcgtgtaaaccaacaacactgataattcattctca 1390
QY 1216 gagaactcttgcaagcattatcagagagatattgtctgggtgttatctgcagttacc 1275
DB |||||||ttgaactgtctaagattgacagacagaataagacacaacacactgtgactatg 1450
1391 gagaactcttgcaagcattatcagagagatattgtctgggtgttatctgcagttacc 1450
QY 1276 tgccttggaaacattttgtcattttgcatgtgacacttatatcagctcgtgagaacgctg 1335
DB |||||||ttgaactgtctaagattgacagacagaataagacacaacacactgtgactatg 1510
1451 tgccttggaaacattttgtcattttgcatgtgacacttatatcagctcgtgagaacgctg 1510
QY 1336 tatgcaatgcaatcattctctcctcgtctgtgcagactgcttaattggaaatatattatc 1395
DB |||||||ttgaactgtctaagattgacagacagaataagacacaacacactgtgactatg 1570
1511 tatgcaatgcaatcattctctcctcgtctgtgcagactgcttaattggaaatatattatc 1570
QY 1396 gtgactggaagccttgacccaagattcgtggaataatacaataagcactgcagctgtgg 1455
DB |||||||ttgaactgtctaagattgacagacagaataagacacaacacactgtgactatg 1630
1571 gtgactggaagccttgacccaagattcgtggaataatacaataagcactgcagctgtgg 1630
QY 1456 atgagaatcatcatlctcagcctgttagaactcttggcaactctgtccacagaagtatca 1515
DB |||||||ttgaactgtctaagattgacagacagaataagacacaacacactgtgactatg 1690
1631 atgagaatcatcatlctcagcctgttagaactcttggcaactctgtccacagaagtatca 1690
QY 1516 gtttactgttaacattctgcaactggaataataacatcgcgactgtgctatcctttaga 1575
DB |||||||ttgaactgtctaagattgacagacagaataagacacaacacactgtgactatg 1750
1691 gtttactgttaacattctgcaactggaataataacatcgcgactgtgctatcctttaga 1750
QY 1576 tgtgtgagactgtgaaatgcaagaacaaattagactttgattctcatcttgattgattg 1635
DB |||||||ttgaactgtctaagattgacagacagaataagacacaacacactgtgactatg 1810
1751 tgtgtgagactgtgaaatgcaagaacaaattagactttgattctcatcttgattgattg

QY	1636	tttaagtgcttcaatccatgagcaataaggaaattttcaaaactactatggcacc	1695
Db	1811	tttatagtgcttcaatccatccatgagcaataaggaaattttcaaaactactatggcacc	1870
QY	1696	aatgagatgcttccctcttcaatcaagaagataagaaagattggagcccaattat	1755
Db	1871	aatgagatgcttccctcttcaatcaagaagataagaaagattggagcccaattat	1930
QY	1756	tcaatgcaattttcttgatgtaattttggccgaattatcaatgatttttccat	1815
Db	1931	tcaatgcaattttcttgatgtaattttggccgaattatcaatgatttttccat	1990
QY	1816	ggaagcctgtttatagtggttcaaaagtgcacataacgagcagcaatgcgaatca	1875
Db	1991	ggaagcctgtttatagtggttcaaaagtgcacataacgagcagcaatgcgaatca	2050
QY	1876	gttaaaaaaagatgataccttgcacaaagttttctttatagtaattacgtatgata	1935
Db	2051	gttaaaaaaagatgataccttgcacaaagttttctttatagtaattacgtatgata	2110
QY	1936	tgctggtatcccatlcttggtagtgaaattcttccactgcttcagtgagaaatacaggt	1995
Db	2111	tgctggtatcccatlcttggtagtgaaattcttccactgcttcagtgagaaatacaggt	2170
QY	1996	accatacccttgggtagtgatattttatcttgcacataacagtgcttgaaccgaatt	2055
Db	2171	accatacccttgggtagtgatattttatcttgcacataacagtgcttgaaccgaatt	2230
QY	2056	ctctatctcttgcacaaagacacatttaagaatgatcatcaggtttgtgataaac	2115
Db	2231	ctctatctcttgcacaaagacacatttaagaatgatcatcaggtttgtgataaac	2290
QY	2116	agacaaagaataactatgagcagcaaaaggtgcgaagaacataatgcccacatcatctgg	2175
Db	2291	agacaaagaataactatgagcagcaaaaggtgcgaagaacataatgcccacatcatctgg	2350
QY	2176	gttgaaatgtgccaactgcagagatgccaactgagttatgaaagcggaacctttcaca	2235
Db	2351	gttgaaatgtgccaactgcagagatgccaactgagttatgaaagcggaacctttcaca	2410
QY	2236	tacccctgtgaaatgctacactgattttcacaataacagagactaatctctatcatga	2292
Db	2411	tacccctgtgaaatgctacactgattttcacaataacagagactaatctctatcatga	2467
RESULT 3			
AA290524			
ID	AA290524 standard; cDNA: 1804 BP.		
XX			
AC	AA290524:		
XX			
DT	05-JUN-2000 (first entry)		
XX			
DE	Human GPCR protein (HGPRP) encoding cDNA (clone ID 2488822).		
XX			
KW	Human; G protein coupled protein receptor; HGPRP; cell proliferation;		
KW	neurological; immune disorder; cytosolic; anti-arteriosclerotic;		
KW	anti-atherosclerotic; hepatotropic; antiinflammatory; virucide; leukemia;		
KW	immunomodulatory; anemia; asthma; gastrointestinal; anti-epileptic;		
XX	anti-Alzheimer's; anti-Parkinsonian; gene therapy; ss.		
OS	Homo sapiens.		
XX			
PN	W0200015793-A2.		
XX			
FD	23-MAR-2000.		
XX			
PF	17-SEP-1999; 99MO-US20958.		
XX			
PR	17-SEP-1998; 98US-0156513.		
XX			
PA	(INCY-) INCYTE PHARM INC.		

xx	Bardman O,	Lal P,	Tang YT,	Corley NC,	Guebler KU,	Gorgone GA;
pi	Baughn MR;					
xx						
DR	WPI: 2000-271432/23.					
xx	P-PDB; AAY57286.					
xx						
PT	Human G protein coupled protein receptor peptides useful for the					
PT	prevention, diagnosis and treatment of cell proliferative, neurological					
PT	and immune disorders -					
xx						
PS	Claim 9; Page 69-70; 71pp; English.					
xx						
CC	The invention provides human G protein coupled protein receptor (HGPRP)					
CC	polypeptides and polynucleotides encoding them. The polypeptides can be					
CC	produced by standard recombinant methodology. The polynucleotides and					
CC	polypeptides may be used in the prevention, treatment and diagnosis of					
CC	diseases associated with their inappropriate expression. Diseases that					
CC	can be treated are cell proliferative disorders (e.g. arteriosclerosis,					
CC	atherosclerosis and hepatitis), cancers (e.g. leukemia, melanomas and					
CC	adenocarcinoma), immune disorders (e.g. anemia, asthma and Crohn's					
CC	disease) and neurological disorders (e.g. epilepsy, Alzheimer's disease					
CC	and Parkinson's disease). The anti-HGPRP antibodies may also be used as					
CC	diagnostic agents for detecting the presence of HGPRP polypeptides in					
CC	samples (e.g. by enzyme linked immunosorbent assay (ELISA)). Sequences					
CC	AA030521-526 represent cDNA fragments encoding the HGPRP polypeptides					
CC	(AAY57283-288).					
SQ	Sequence 1804 BP; 563 A; 318 C; 327 G; 596 T; 0 other;					
	Query Match	46.2%;	Score 1656.8;	DB 21;	Length 1804;	
	Best Local Similarity	99.2%;	Pred. No. 0;			
	Matches 1675;	Conservative 0;	Mismatches 12;	Indels 1;	Gaps 1;	
OY	105f tcttcagcctagaaggattgaattcaattccaatccaaagaagtgttagaccttt 1116					
Db	105 tttagcacgctagaaggattgaattcaattccaatccaaagaagtgttagaccttt 164					
OY	1117 atgaatctctccacatatattttaagaattccagtaactgtgggtatgcaccacatgtt 1176					
Db	165 gtgaatctctccacatatattttaagaattccagtaactgtgggtatgcaccacatgtt 224					
OY	1177 gcagcgtgtaaaccacaacactgtatgaattcatctctagaagatctcttgccaagcat 1236					
Db	225 gcgagctgtaaacccaacacgtatggaattcatctctagaagatctcttgccaagcat 284					
OY	1237 attcagaagatatttgtctgggtgttatctcgcaattacctgctttggaaacattttgtc 1296					
Db	285 attcagagagtatattgtctgggtgttatctcgcaattacctgctttggaaacattttgtc 344					
OY	1297 atttcatacgacccttatatcagrgctctagaaccaagcgtgatgccatltcaaatcttct 1356					
Db	345 atttcatacgacccttaatacagrgctctagaaccaagcgtgatgccatltcaaatcttct 404					
OY	1357 ctctcgtgtgcccagactgtcttaatgyggaatataatttatctgtgatcgagggtcttgactba 1416					
Db	405 ctctcgtgtgcccagactgtcttaatgyggaatataatttatctgtgatcgagggtcttgactba 464					
OY	1417 aaatttcgttgagaataacaataagaatgcagcgtgtggaatgagagaagtaactattgtcaa 1476					
Db	465 aaatttcgttgagaataacaataagaatgcagcgtgtggaatgagagaagtaactattgtcaa 524					
OY	1477 ctgtgtagagctcttggccattctctgccccagaagaafataagtttaactctttaacatttctg 1536					
Db	525 ctgtgtagagctcttggccattctctgccccagaagaafataagtttaactctttaacatttctg 584					
OY	1537 acatttgaaaaataacatctgcattgtctatccttttagatgtgtgagaccttgaaaatgac 1596					
Db	585 acatttgaaaaataacatctgcattgtctatccttttagatgtgtgagaccttgaaaatgac 644					
OY	1597 agaaacaattaaagttctgattctcaatttggaattactgatttataagvtggtcttcaatca 1656					

Db	645	agaacaatlaacagltcgtcattccatttcgtattacacgtgtttatagtcgttcaatcca	704
OY	1657	ttgagcaataaaggaaattttccaaaactactatgycaccaaigtgagtagtcttcctct	1716
Db	705	ttgagcaataaaggaaattttccaaaactactatgycaccaaigtgagtagtcttcctct	764
OY	1717	cattcaagaagtaagaagaattttgagcccaatttatccgttgcaattttcttgt	1776
Db	765	cattcaagaagtaagaagaattttgagcccaatttatccgttgcaattttcttgt	824
OY	1777	attaatttggccgcatttatcatcatatggttttccatagtagaagcagtgtttatagtg	1836
Db	825	attaatttggccgcatttatcatcatatggttttccatagtagaagcagtgtttatagtg	884
OY	1837	catcaaaagtcacataacagcaactgaaatacggaaatcaagttaaaaaagagatgcct	1896
Db	885	catcaaaagtcacataacagcaactgaaatacggaaatcaagttaaaaaagagatgcct	944
OY	1897	gccaaaagtttttcttcttatagatattacacgacattatgctcgtgatacccaatttta	1956
Db	945	gccaaaagtttttcttcttatagatattacacgacattatgctcgtgatacccaatttta	1004
OY	1957	gtgaaattcttcttcactgcgtccctcagtagaataccaggtacataaactccttggtag	2016
Db	1005	gtgaaattcttcttcactgcgtccctcagtagaataccaggtacataaactccttggtag	1064
OY	2017	atttttatctgcgccattaacagtgctttggaaccacaattcctctatacctctgcacaa	2076
Db	1065	atttttatctgcgccattaacagtgctttggaaccacaattcctctatacctctgcacaa	1124
OY	2077	ccatttaagaagaagatcagtcacgtgtttgttatcaataagaacaaagaatctatgac	2136
Db	1125	ccatttaagaagaagatcagtcacgtgtttgttatcaataagaacaaagaatctatgac	1184
OY	2137	agcaaaagtcagaaaaaataatgctccatcatatcatcttgggttggaatgtrgcactcag	2196
Db	1185	agcaaaagtcagaaaaaataatgctccatcatatcatcttgggttggaatgtrgcactcag	1244
OY	2197	gagatgcacactggtgtaataagaacggagacctttacaataccctgtgaaatgtrcactg	2256
Db	1245	gagatgcacactggtgtaataagaacggagacctttacaataccctgtgaaatgtrcactg	1304
OY	2257	attctcaatacaagagagacccaattccctatactatgctgcacctgaaattactcttcg	2316
Db	1305	attctcaatacaagagagacccaattccctatactatgctgcacctgaaattactcttcg	1364
OY	2317	cagaagaactctgtagggtgcttcctatgtagggaatttactgtagtaaaatgaataccaa	2376
Db	1365	cagaagaactctgtagggtgcttcctatgtagggaatttactgtagtaaaatgaataccaa	1423
OY	2377	aattaatttataataataagctaaagataaataatttacaagaagcagtgaagaaaaataaa	2436
Db	1424	aattaatttataataataagctaaagataaataatttacaagaagcagtgaagaaaaataaa	1483
OY	2437	tgcataatgcctcttaacaagaaggaatataatatacaataatataatattatagtagca	2496
Db	1484	tgcataatgcctcttaacaagaaggaatataatatacaataatataatattatagtagca	1543
OY	2497	ttttgcataaagaataaagaagaactactctcgaatacatcatcatctttctcaacg	2556
Db	1544	ttttgcataaagaataaagaagaactactctcgaatacatcatcatctttctcaacg	1603
OY	2557	catttatgtagtagtaaccactactatgtagcatagcatgtaataatagctcgtgaagtaga	2616
Db	1604	catttatgtagtagtaaccactactatgtagcatagcatgtaataatagctcgtgaagtaga	1663
OY	2617	gtcgaagaactttcaatctgttagatagtggtttaatgacaagaagctatacaaaatgcac	2676
Db	1664	gtcgaagaactttcaatctgttagatagtggtttaatgacaagaagctatacaaaatgcac	1723
OY	2677	tgcagttcccaagttaagaagaagccttaaccgttaatgtagcatcgaacgaagatcatagc	2736
Db	1724	tgcagttcccaagttaagaagaagccttaaccgttaatgtagcatcgaacgaagatcatagc	1783

QY	2737	actitaa	2744
Db	1784	actitaa	1791
RESULT	4		
ID	AAD06507		
AC	AAD06507		
XX			
DF	10-AUG-2001	(first entry)	
XX			
DE	Human CON222 G protein-coupled receptor DNA.		
XX			
KM	Human; G protein-coupled receptor; GPCR; CON222 protein; schizophrenia; neuroleptic; nootropic; neuroprotective; bipolar disease; psychotropic;		
KM	neurological disorder; psychiatric disease; neurosis; anxiety; neuritis;		
KM	attention deficit hyperactivity disorder; neuroathenia; senile dementia;		
KM	affective disorder; neuropathy; Alzheimer's disease; Parkinson's disease;		
KM	depression; migraine; genetic screening; ds.		
OS			
XX	Homo sapiens.		
FH			
FT	Key	Location/Qualifiers	
FT	CDS	1..1191	
FT		/*tag= a	
XX		/product= "Human G protein-coupled receptor protein"	
XX			
XX	WO200131014-A2.		
XX			
PD	03-MAY-2001.		
XX			
PF	27-OCT-2000; 2000WO-US29601.		
XX			
XX	27-OCT-1999; 99US-0427653.		
PR	27-OCT-1999; 99US-0427859.		
PR	27-OCT-1999; 99US-0428020.		
PR	27-OCT-1999; 99US-0428114.		
PR	28-OCT-1999; 99US-0429517.		
PR	28-OCT-1999; 99US-0429555.		
PR	28-OCT-1999; 99US-0429676.		
PR	28-OCT-1999; 99US-0429695.		
PR	03-DEC-1999; 99US-0454399.		
PR	12-JAN-2000; 2000US-0481794.		
XX			
PA	(PHAA) PHARMACIA & UPJOHN CO.		
XX			
PI	Vogel I, Wood LS, Merchant K;		
XX			
DR	WPI; 2001-328653/34.		
DR	P-PSDB; AAE02498.		
XX			
PT	Seven transmembrane receptor polypeptides and polynucleotides, useful		
PT	for treating neuroleptic or psychiatric disorders, e.g. schizophrenia,		
PT	as well as for identifying compounds useful for treating schizophrenia		
XX			
PS	Claim 22; Page 15-16; 215pp; English.		
XX			
CC	The invention relates to human G protein-coupled receptor (GPCR) and		
CC	their corresponding DNA molecules. GPCR is also referred as seven		
CC	transmembrane receptor. G protein-coupled receptor protein is useful for		
CC	treating neurological disorder, particularly schizophrenia. GPCR protein		
CC	is also useful for identifying compounds useful for treating		
CC	schizophrenia. These compounds are also useful for treating other		
CC	neurological and psychiatric diseases, e.g. depression, anxiety, bipolar		
CC	disease, affective disorders, attention deficit hyperactivity disorder/		
CC	attention deficit disorder, epilepsy, neuritis, neuroathenia, neuropathy		
CC	neurosis, Alzheimer's disease, Parkinson's disease, migraine and senile		
CC	dementia. The invention also provides genetic screening procedures that		
CC	entail analysing a person's genome with respect to GPCR. The vectors are		

CC useful for the recombinant production of the GPCR's. The present DNA
 CC sequence encodes human CON222 G protein-coupled receptor (GPCR) protein.
 XX
 SQ Sequence 1191 BP; 340 A; 229 C; 226 G; 396 T; 0 other;

Query Match 33.2%; Score 1189.4; DB 22; Length 1191;
 Best Local Similarity 99.9%; Pred. No. 6,9e-268;
 Matches 1190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1102 atgttagaccttatagtaatctctctcaacatattttaagaattccagctactgagg 1161
 |||||||
 Db 1 atgttagacctctgtgaatctctctcaacatattttaagaattccagctactgagg 60
 QY 1162 tatgaccacatgtctgcagctgtgtaaaccaaacctgatagtgaatttcattctctagaat 1221
 |||||||
 Db 61 tatgaccacatgtctgcagctgtgtaaaccaaacctgatagtgaatttcattctctagaat 120
 QY 1222 cctctggcaagcatattccagagagatttgtctgggttgatctgcagttactgcttt 1281
 |||||||
 Db 121 cctctggcaagcatattccagagagatttgtctgggttgatctgcagttactgcttt 180
 QY 1282 ggaacattttgtcatttgcatacgcaacctatatacagtlctgagaaacagctgtatgcc 1341
 |||||||
 Db 181 ggaacattttgtcatttgcatacgcaacctatatacagtlctgagaaacagctgtatgcc 240
 QY 1342 atgttcaatcttctctctctgtctgtgcagctgttaatgggaatatattttcgtatc 1401
 |||||||
 Db 241 atgttcaatcttctctctctgtctgtgcagctgttaatgggaatatattttcgtatc 300
 QY 1402 ggaagcttaccctaagaagttctgtgagaaatacaataagcagctgcagctgtgagtga 1461
 |||||||
 Db 301 ggaagcttaccctaagaagttctgtgagaaatacaataagcagctgcagctgtgagtga 360
 QY 1462 agtactcatgtcagctgtgtagaactcttggcattctgttccagaagaagtlcagttta 1521
 |||||||
 Db 361 agtactcatgtcagctgtgtagaactcttggcattctgttccagaagaagtlcagttta 420
 QY 1522 ctgttaacatttctggaattggaaaaatacatctgcattgtctactcttttaagtgtg 1581
 |||||||
 Db 421 ctgttaacatttctggaattggaaaaatacatctgcattgtctactcttttaagtgtg 480
 QY 1582 agacctggaanaatgcagaaacattacagtttctgattctcatttgattactggtttta 1641
 |||||||
 Db 481 agacctggaanaatgcagaaacattacagtttctgattctcatttgattactggtttta 540
 QY 1642 gtggcttcatctccattgagcaataaagaattttccaaaaactactatgscacaatgga 1701
 |||||||
 Db 541 gtggcttcatctccattgagcaataaagaattttccaaaaactactatgscacaatgga 600
 QY 1702 gtatgcttccctcttcttcaagaagatacagaagaatttggagcccgattattcagtg 1761
 |||||||
 Db 601 gtatgcttccctcttcttcaagaagatacagaagaatttggagcccgattattcagtg 660
 QY 1762 gcaatttctgtgtataatttggccgcatattcatcattagtttttctcattggaagc 1821
 |||||||
 Db 661 gcaatttctgtgtataatttggccgcatattcatcattagtttttctcattggaagc 720
 QY 1822 atgtttatagtggttcataaagtgcataacagcaactgtaatacaggaatcaagtttaa 1881
 |||||||
 Db 721 atgtttatagtggttcataaagtgcataacagcaactgtaatacaggaatcaagtttaa 780
 QY 1882 aaagaagatgctctggcaaacgtttttctttagtatttactagtgactgtgcgg 1941
 |||||||
 Db 781 aaagaagatgctctggcaaacgtttttctttagtatttactagtgactgtgcgg 840
 QY 1942 ataccattttttagtgaattcttcttcaactgtctcaggtagaataacaggtacata 2001
 |||||||
 Db 841 ataccattttttagtgaattcttcttcaactgtctcaggtagaataacaggtacata 900
 QY 2002 acctctgggtagtgattttattctgcacataacagtgcttgaaccacattctctat 2061
 |||||||
 Db 901 acctctgggtagtgattttattctgcacataacagtgcttgaaccacattctctat 960

QY 2062 acctgaccacaagaccatttaagaatagtatcatcggtttgttataactacagaa 2121
 |||||||
 Db 961 acctgaccacaagaccatttaagaatagtatcatcggtttgttataactacagaa 1020
 QY 2122 agaaatctatgagcagcaaaaggtcagaanaacatatgtccatcatcatcattcgtgtgaa 2181
 |||||||
 Db 1021 agaaatctatgagcagcaaaaggtcagaanaacatatgtccatcatcatcattcgtgtgaa 1080
 QY 2182 atgtggccactgacagagatgcccacctgagtttaatagaagccgagacctttacataccc 2241
 |||||||
 Db 1081 atgtggccactgacagagatgcccacctgagtttaatagaagccgagacctttacataccc 1140
 QY 2242 tgttaaatgtacagtatttccatcaacagagactcaattccttattga 2292
 |||||||
 Db 1141 tgtgaatgtacagtatttccatcaacagagactcaattccttattga 1191

RESULT 5
 AA199557
 ID AA199557 standard; cDNA; 1089 BP.
 XX
 AC AA199557;
 XX
 DT 04-JAN-2002 (first entry)
 XX
 DE Human expressed polynucleotide SEQ ID NO 20.
 XX
 XX Human; nootropic; neuroprotective; cytostatic; dermatological; virocidic;
 KW immunosuppressive; anti-inflammatory; anti-HIV; antibacterial; vulnerary;
 KW antiparkinsonian; antischizoid; antianemic; antiallergic; cancer;
 KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
 KW antiallergic; antidiabetic; antileuc; anticonvulsant; antifungal;
 KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; nephrotropic; gene therapy; vaccine;
 ss.
 XX
 OS Homo sapiens.
 XX
 PN W0200155387-A1.
 XX
 PD 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01310.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.

PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233065.
PR 14-SEP-2000; 2000US-0233066.
PR 14-SEP-2000; 2000US-0233067.
PR 14-SEP-2000; 2000US-0233068.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 25-SEP-2000; 2000US-0234999.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246539.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.

PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250161.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251038.
PR 05-DEC-2000; 2000US-0251989.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-451937/48.

P-PSDB; ABB04105.

Isolated polypeptide for treating, preventing and/or prognosing disorders related to the musculoskeletal system including musculoskeletal cancers and also for testing and detection e.g. diagnosis -

Claim 1; SEQ ID NO 1029; 781bp + Sequence Listing; English.

The invention relates to novel genes (AAL34669-AAL37666) and proteins (AAB03087-AAB04109) associated with the musculoskeletal system useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.

CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 760 BP; 235 A; 144 C; 148 G; 226 T; 7 other;

Query Match 18.7%; Score 668.8; DB 22; Length 760;
Best Local Similarity 96.7%; Pred. No. 2,3e-146;
Matches 712; Conservative 0; Mismatches 18; Indels 6; Gaps 3;

QY 852 gaggaanaaataataactactaaatgaatgaatcttgcaccttcagaactcga 911
DB 1 gaggaanaaataataactactaaatgaatgaatcttgcaccttcagaactcga 60
QY 912 tgaattgattggaagtaataagatgaatgaatcttcacccggtatattcaagacct 971
DB 61 tgaattgattggaagtaataagatgaatgaatcttcacccggtatattcaagacct 120
QY 972 gaaggaagctcacaattgaaatcttcctataatccaatccaagaatattcaagaacca 1031
DB 121 gaaggaagctcacaattgaaatcttcctataatccaatccaagaatattcaagaacca 180
QY 1032 attgattatctgtccaactcaagtcctcgaagcctgaaggaattgaaattcaatat 1091

```

Db 181 attgattattctgttcaaccctgaagctctc---cctagaaggagattgaattccaatat 237
Oy 1092 ccaacaagaagatgttttagacctctatgaatcctctctacatatatttaagaatcca 1151
Db 238 ccaataaagaagatgttttagacctctctatgaatcctctctacatatatttaagaatcca 297
Oy 1152 gctactgtggaatgacacacatgttcgagctgtgaacacacactgttgtaattctc 1211
Db 298 gtaactgtggaatgacacacatgttcgagctgtgaacacacactgttgtaattctc 357
Oy 1212 tctagaagatctctctgcaacacatattcagagaagatttctgtgtgtatctgcagt 1271
Db 358 tctagaagatctctctgcaacacatattcagagaagatttctgtgtgtatctgcagt 417
Oy 1272 tactgtcttggaaacatttctgtcatattgcatgcgacctatatacagctcgtgaacaa 1331
Db 418 tactgtcttggaaacatttctgtcatattgcatgcgacctatatacagctcgtgaacaa 477
Oy 1332 gctgtatgcatgtcaatcattctctctgtctgtgcgacctgaatgggaatatatt 1391
Db 478 gctgtatgcatgtcaatcattctctctgtctgtgcgacctgaatgggaatatatt 537
Oy 1392 attcgtatcggaagcttgcacctaagaatttcgtgagaatacaataagacatgcgcagc- 1450
Db 538 attcgtatcggaagcttgcacctaagaatttcgtgagaatacaataagacatgcgcagc 597
Oy 1451 tctgtatgagaagatcattcattgtcagctgtga--ggaactcttgccactctgtccacaga 1508
Db 598 tctgtatgagaagatcattcattgtcagctgtga--ggaactcttgccactctgtccacaga 657
Oy 1509 agtaccagtttactcgttaacattctgcacattctggaacaaatacactcattgtctatcc 1568
Db 658 agtaccagtttactcgttaacattctgcacattctggaacaaatacactcattgtctatcc 717
Oy 1569 tttagatgtgtgaga 1584
Db 718 nttagaaggaggaggaga 733

```

RESULT 7
 AAA44932
 ID AAA44932 standard; CDNA; 636 BP.
 XX
 AC AAA44932;
 XX
 DT 21-AUG-2000 (first entry)
 XX
 DE Human secreted expressed sequence tag SEQ ID NO:1507.
 XX
 KW Human: mouse; chicken; rat; secreted expressed sequence tag; SEST; expressed sequence tag; EST; probe; chemotactic; proliferative; immunomodulatory; haematopoietic; chemokine; analgesic; haemostatic; thrombolytic; antiinflammatory; cytostatic; antidiabetic; antifungal; antiviral; antidiabetic; antisthmatic; vulnary; antiparkinsonian; antitumor; osteoprotective; neuroprotective; neurotropic; antiparasitic; cerebroprotective; anticonvulsant; antidepressant; gene therapy; vaccine; autoimmune disorder; multiple sclerosis; allergic condition; insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer; lymphoid cell deficiency; burn; osteoporosis; osteoarthritis; central nervous system disorder; Alzheimer's disease; stroke; Parkinson's disease; Huntington's disease; coagulation disorder; haemophilia; thrombosis; inflammatory disorder; Crohn's disease; tumour; infection; depression; psoriasis; ss.
 KW
 OS Homo sapiens.
 OS
 PN WO200021991-A1.
 XX
 PD 20-APR-2000.
 XX
 PF 15-OCT-1999; 99WO-US24206.
 XX

```

PR 15-OCT-1998; 98US-0104436.
XX
PA (GEMT ) GENETICS INST INC.
XX
PI Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;
XX Merberg D, Treacy M, Bowman MR;
XX
DR WPI: 2000-317938/27.
XX
XX Isolated polynucleotides, and encoded proteins, comprising secreted
PT expressed sequence tags (SESTs), useful for treating various disorders
PT such as autoimmune, infectious, and central nervous system disorders -
XX
XX Claim 1; Page 593; 803pp; English.
XX
XX AAA43426 to AAA5925 represent specifically claimed secreted expressed
CC sequence tags (SESTs), isolated from human, mouse, chicken and rat
CC tissue sources. The SESTs can have a range of activities depending on
CC the tissues they were isolated from. The activities include:
CC chemotactic; proliferative; immunomodulatory; haematopoietic;
CC chemokine; analgesic; haemostatic; thrombolytic; antidiabetic;
CC cyostatic; antibacterial; antifungal; antiviral; antidiabetic;
CC antisthmatic; vulnary; antitumor; osteoprotective;
CC neurotropic; antiparkinsonian; antiparasitic; cerebroprotective;
CC anticonvulsant; and antidepressant. The SESTs can be used for gene
CC therapy and in vaccines. The SESTs are useful as probes for the
CC identification and isolation of full-length cDNAs and genomic DNA
CC molecules which correspond to the SESTs. Proteins encoded by the SESTs
CC are useful in assays for determining biological activity and raising
CC antibodies. They may be useful for treatment of autoimmune disorders
CC (multiple sclerosis, insulin dependent diabetes), allergic conditions
CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,
CC osteoporosis, osteoarthritis, central nervous system disorders
CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation
CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's
CC disease), tumours, bacterial, fungal or viral infections, depression and
CC psoriasis. AAA5926 to AAA5931 represent linker variants which are given
CC in the exemplification of the present invention.
XX
XX Sequence 636 BP; 217 A; 115 C; 112 G; 186 T; 6 other:
SQ

```

Query Match 16.2%; Score 580.4; DB 21; Length 636;
 Best Local Similarity 98.7%; Pred. No. 9.3e-126;
 Matches 603; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

```

Oy 2016 gattttattctgcccattacagatgcttgaacccaattctctactctgaccacaag 2075
Db 22 gattttattctgcccattacagatgcttgaacccaattctctactctgaccacaag 81
Oy 2076 accatttaagaagaatgattcattcggttttggatataactacagaacagaagaatct-attg 2134
Db 82 accatttaagaagaatgattcattcggttttggatataactacagaacagaagaatctcnnhgg 141
Oy 2135 acaagaagaagtcagaacacatactgctcatcattcattcgggttggaatgtgacactgc 2194
Db 142 acaagaagaagtcagaacacacatactgctcatcattcattcgggttggaatgtgacactgc 201
Oy 2195 aggaagatgcacactgagtaagtcgaagccggacatttcaacataccctgtgaaagtcac 2254
Db 202 aggaagatgcacactgagtaagtcgaagccggacatttcaacataccctgtgaaagtcac 261
Oy 2255 tgatttccaatcaacagagactcaattcctatcaatgactgactggaattcattctt 2314
Db 262 tgatttccaatcaacagagactcaattcctatcaatgactgactggaattcattctt 321
Oy 2315 cgcagaagaatctgagggtgcttcacatgagagattctgtaagaagaatgaataacac 2374
Db 322 cgcagaagaatctgagggtgcttcacatgagagattctgtaagaagaatgaataacac 380
Oy 2375 aaaaatttaataataagtcagaataatatttcaagaagacatgaggaataataa 2434
Db 381 aaaaatttaataataagtcagaataatatttcaagaagacatgaggaataataa 440

```

QY 2435 aatgactaatgctcttaacaaagggaatattatatacaataatgtatatattagtaga 2494
|||||
Db 441 aatgactaatgctcttacttaacaaagggaatattatatacaataatgtatatattagtaga 500
QY 2495 catttgcataagaataatgaagaatctacttaagtaacatcaatcattttcttaaca 2554
|||||
Db 501 catttgcataagaataatgaagaatctacttaagtaacatcattcattttcttaaca 560
QY 2555 tgcattattatagaccactactatgtgcataagcatttgcataatagtccctggaagttaga 2614
|||||
Db 561 tgcattattatagaccactactactatgtgcataagcatttgcataatagtccctggaagttaga 620
QY 2615 cagtcagagaac 2625
|||||
Db 621 cagtcagagaac 631
RESULT 8
AA199584
ID AA199584 standard; cDNA; 530 BP.
XX
AC AA199584;
XX
DT 04-JAN-2002 (first entry)
XX
DE Human expressed polynucleotide SEQ ID NO 47.
XX
KW Human; nootropic; neuroprotective; cytosolic; dermatological; virucide;
KW immunosuppressive; antihistaminic; anti-HIV; antibacterial; vulnerrary;
KW antiparkinsonian; antischistosomal; antitubercular; cancer;
KW antileukemic; hepatotoxic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine;
KW ss.
XX
XX Homo sapiens.
OS
XX
PN WO20015387-A1.
XX
PD 02-AUG-2001.
XX
PE 17-JAN-2001; 2001WO-US01310.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-MAR-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218280.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227909.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 06-SEP-2000; 2000US-0231242.
PR 06-SEP-2000; 2000US-0231243.
PR 06-SEP-2000; 2000US-0231243.
PR 06-SEP-2000; 2000US-0231244.
PR 06-SEP-2000; 2000US-0231413.
PR 06-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 01-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.

PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI: 2001-465573/50.
DR P-PSDB: AAM99972.
XX
XX
PT Isolated digestive system associated polypeptide for treating,
PT preventing and/or prognosing disorders related to the digestive system
PT including digestive system cancers and also for testing and detection
PT e.g. diagnosis -
XX
XX
PS Claim 1; SEQ ID NO 47; 509bp + Sequence Listing; English.
XX
CC The invention relates to novel genes (AA199548-AA199604) and proteins
CC (AAM99936-AA99984) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases such as cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 530 BP; 158 A; 105 C; 95 G; 171 T; 1 other;

Query Match 14.7%; Score 528.6; DB 22; Length 530;

Best Local Similarity 99.8%; Pred. No. 1.1e-113;
Matches 528; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1744 gccacagattatcagtcggaattttctctgtgtatattgtgcgcattatcatcata 1803
DB 2 gccacagattatcagtcggaattttctctgtgtatattgtgcgcattatcatcata 61
QY 1804 gtttttccatggaagcagttttatagtcgttcataaagtgccataagcagcagaa 1863
DB 62 gtttttccatggaagcagttttatagtcgttcataaagtgccataagcagcagaa 121
QY 1864 atacggaatcaagtaaaagagatgatccttgcacaaagttttcttatagta 1923
DB 122 atacggaatcaagtaaaagagatgatccttgcacaaagttttcttatagta 181
QY 1924 acggaatcaagtaagtcagataccattttgtagtgaaatttcttactgtctcagta 1983
DB 182 actgatgatatttgctgataccattttgtagtgaaatttcttactgtctcagta 241
QY 1984 gaaataccaggtaccataacctctgggtagtgatttttcttgcacataagtgct 2043
DB 242 gaaataccaggtaccataacctctgggtagtgatttttcttgcacataagtgct 301
QY 2044 ttgaaccacattctctactctgacacaaagacatttaagaatgatcatcgctt 2103
DB 302 ttgaaccacattctctactctgacacaaagacatttaagaatgatcatcgctt 361
QY 2104 tggataactacagacaagaanaatctatgacagcaagtcagaaacatatgtcca 2163
DB 362 tggataactacagacaagaanaatctatgacagcaagtcagaaacatatgtcca 421
QY 2164 tcatctcctgggtggaagtgtgcccactgcagaagttgcacactggttaataagcgg 2223
DB 422 tcatctcctgggtggaagtgtgcccactgcagaagttgcacactggttaataagcgg 481
QY 2224 gacctttcacataccctcgtgaaatgtcactgatttctcaatcaagca 2272
DB 482 gacctttcacataccctcgtgaaatgtcactgatttctcaatcaagca 530
RESULT 9
AAL35644
ID AAL35644 standard; cDNA; 530 BP.
XX
AC AAL35644;
XX
DT 08-JAN-2002 (first entry)
XX
DE Human musculoskeletal system related polynucleotide SEQ ID NO 986.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
KW vulnerrary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein;
KW musculoskeletal system; ss.
XX
OS Homo sapiens.
XX
PN W0200155367-A1.
XX
PD 02-NOV-2001.
XX
PF 17-JAN-2001; 2001WO-0501338.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.

CC protein or gene therapy. The genes are isolated from a range of human
CC tissues disclosed in the specification. The nucleic acids, proteins,
CC antibodies and (ant)agonists are useful in the diagnosis, treatment
CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and
CC other cancers of the adrenal gland, bone, bone marrow, breast,
CC gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC (c) cardiovascular disorders such as myocardial ischaemia; (d) wound
CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
CC and (f) infectious diseases such as viral, bacterial, fungal and
CC parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 530 BP; 158 A; 105 C; 95 G; 171 T; 1 other;

Query Match 14.7%; Score 528.6; DB 22; Length 530;
Best Local Similarity 99.8%; Pred. No. 1.1e-113;
Matches 528; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1744 gccacagattatcacagtggaattttcttgatataattggcgcatattatcata 1803
|||
DB 2 gccacagattatcacagtggaattttcttgatataattggcgcatattatcata 61
QY 1804 gtttttccataggaagcatgtttatagtggtcatcaagtgcaacacatgaa 1863
|||
DB 62 gtttttccataggaagcatgtttatagtggtcatcaagtgcaacacatgaa 121
QY 1864 atacggagataagtaaaagaagaatgatccttgccaaagtttttccatagtatt 1933
|||
DB 122 atacggagataagtaaaagaagaagaatgatccttgccaaagtttttccatagtatt 181
QY 1924 actgatcatattatgctgataccattttgtatgtaaatcttccactgacatgata 1983
|||
DB 182 acgatgcatattatgctgataccattttgtatgtaaatcttccactgacatgata 241
QY 1984 gaaatacaggataacataacccttgggtagtgattttatctgcccataacagtgct 2043
|||
DB 242 gaaatacaggataacataacccttgggtagtgattttatctgcccataacagtgct 301
QY 2044 ttgaacccaattctctactctgacacaaagacatttaagaatgatcatcggtt 2103
|||
DB 302 ttgaacccaattctctactctgacacaaagacatttaagaatgatcatcggtt 361
QY 2104 ttgtataactacagacaagaataatctatgacagacaagtgacagaataacatgctcca 2163
|||
DB 362 ttgtataactacagacaagaataatctatgacagacaagtgacagaataacatgctcca 421
QY 2164 tcattcatctggttggtaaatgtggccactgcaaggagaatgcccactgagttatgaagcgg 2223
|||
DB 422 tcattcatctggttggtaaatgtggccactgcaaggagaatgcccactgagttatgaagcgg 481
QY 2224 gaccttttcaataccctgtgaaatgtaactgatttctcaatacaga 2272
|||
DB 482 gaccttttcaataccctgtgaaatgtaactgatttctcaatacaga 530

RESULT 10

ABA06471
ID ABA06471 standard; cDNA: 530 BP.

XX ABA06471;

XX 10-JAN-2002 (first entry)

XX Human cDNA SEQ ID NO: 137.

XX Human; gene therapy; neural disorder; immune system disorder;

KW muscular disorder; reproductive disorder; gastrointestinal disorder;

KW pulmonary disorder; cardiovascular disorder; renal disorder;

KW proliferative disorder; inflammation; ss.

XX Homo sapiens.

XX WO200154474-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-0501349.

XX 31-JAN-2000; 2000US-179065P.

XX 04-FEB-2000; 2000US-180628P.

XX 24-FEB-2000; 2000US-184564P.

XX 02-MAR-2000; 2000US-186350P.

XX 16-MAR-2000; 2000US-189874P.

XX 17-MAR-2000; 2000US-190076P.

XX 18-APR-2000; 2000US-198123P.

XX 19-MAY-2000; 2000US-205515P.

XX 07-JUN-2000; 2000US-209467P.

XX 28-JUN-2000; 2000US-214886P.

XX 30-JUN-2000; 2000US-215135P.

XX 07-JUL-2000; 2000US-216647P.

XX 07-JUL-2000; 2000US-216880P.

XX 11-JUL-2000; 2000US-217487P.

XX 11-JUL-2000; 2000US-217496P.

XX 14-JUL-2000; 2000US-218290P.

XX 26-JUL-2000; 2000US-220963P.

XX 26-JUL-2000; 2000US-220964P.

XX 14-AUG-2000; 2000US-224518P.

XX 14-AUG-2000; 2000US-224519P.

XX 14-AUG-2000; 2000US-225213P.

XX 14-AUG-2000; 2000US-225214P.

XX 14-AUG-2000; 2000US-225266P.

XX 14-AUG-2000; 2000US-225267P.

XX 14-AUG-2000; 2000US-225268P.

XX 14-AUG-2000; 2000US-225270P.

XX 14-AUG-2000; 2000US-225447P.

XX 14-AUG-2000; 2000US-225757P.

XX 14-AUG-2000; 2000US-225758P.

XX 14-AUG-2000; 2000US-225759P.

XX 18-AUG-2000; 2000US-226279P.

XX 22-AUG-2000; 2000US-226681P.

XX 22-AUG-2000; 2000US-226686P.

XX 22-AUG-2000; 2000US-227182P.

XX 23-AUG-2000; 2000US-227009P.

XX 30-AUG-2000; 2000US-228924P.

XX 01-SEP-2000; 2000US-229287P.

XX 01-SEP-2000; 2000US-229343P.

XX 01-SEP-2000; 2000US-229344P.

XX 01-SEP-2000; 2000US-229345P.

XX 05-SEP-2000; 2000US-229509P.

XX 05-SEP-2000; 2000US-229513P.

XX 06-SEP-2000; 2000US-230437P.

XX 06-SEP-2000; 2000US-230438P.

XX 08-SEP-2000; 2000US-231242P.

XX 08-SEP-2000; 2000US-231243P.

XX 08-SEP-2000; 2000US-231244P.

XX 08-SEP-2000; 2000US-231413P.

XX 08-SEP-2000; 2000US-231414P.

XX 08-SEP-2000; 2000US-232080P.

XX 08-SEP-2000; 2000US-232081P.

XX 12-SEP-2000; 2000US-231968P.

XX 14-SEP-2000; 2000US-232397P.

XX 14-SEP-2000; 2000US-232398P.

XX 14-SEP-2000; 2000US-232399P.

XX 14-SEP-2000; 2000US-232400P.

XX 14-SEP-2000; 2000US-232401P.

XX 14-SEP-2000; 2000US-233063P.

XX 14-SEP-2000; 2000US-233064P.

XX 14-SEP-2000; 2000US-233065P.

XX 21-SEP-2000; 2000US-234223P.

XX 21-SEP-2000; 2000US-234274P.

PR 25-SEP-2000; 2000US-234997P.
PR 25-SEP-2000; 2000US-234998P.
PR 26-SEP-2000; 2000US-235484P.
PR 27-SEP-2000; 2000US-235834P.
PR 27-SEP-2000; 2000US-235836P.
PR 29-SEP-2000; 2000US-236327P.
PR 29-SEP-2000; 2000US-236367P.
PR 29-SEP-2000; 2000US-236368P.
PR 29-SEP-2000; 2000US-236369P.
PR 29-SEP-2000; 2000US-236370P.
PR 02-OCT-2000; 2000US-236802P.
PR 02-OCT-2000; 2000US-237037P.
PR 02-OCT-2000; 2000US-237038P.
PR 02-OCT-2000; 2000US-237039P.
PR 02-OCT-2000; 2000US-237040P.
PR 13-OCT-2000; 2000US-239935P.
PR 13-OCT-2000; 2000US-239937P.
PR 20-OCT-2000; 2000US-240960P.
PR 20-OCT-2000; 2000US-241221P.
PR 20-OCT-2000; 2000US-241785P.
PR 20-OCT-2000; 2000US-241786P.
PR 20-OCT-2000; 2000US-241787P.
PR 20-OCT-2000; 2000US-241808P.
PR 20-OCT-2000; 2000US-241809P.
PR 20-OCT-2000; 2000US-241826P.
PR 01-NOV-2000; 2000US-244617P.
PR 08-NOV-2000; 2000US-246474P.
PR 08-NOV-2000; 2000US-246475P.
PR 08-NOV-2000; 2000US-246476P.
PR 08-NOV-2000; 2000US-246477P.
PR 08-NOV-2000; 2000US-246478P.
PR 08-NOV-2000; 2000US-246523P.
PR 08-NOV-2000; 2000US-246524P.
PR 08-NOV-2000; 2000US-246525P.
PR 08-NOV-2000; 2000US-246526P.
PR 08-NOV-2000; 2000US-246527P.
PR 08-NOV-2000; 2000US-246528P.
PR 08-NOV-2000; 2000US-246529P.
PR 08-NOV-2000; 2000US-246532P.
PR 08-NOV-2000; 2000US-246609P.
PR 08-NOV-2000; 2000US-246610P.
PR 08-NOV-2000; 2000US-246611P.
PR 08-NOV-2000; 2000US-246613P.
PR 17-NOV-2000; 2000US-249207P.
PR 17-NOV-2000; 2000US-249208P.
PR 17-NOV-2000; 2000US-249209P.
PR 17-NOV-2000; 2000US-249210P.
PR 17-NOV-2000; 2000US-249211P.
PR 17-NOV-2000; 2000US-249212P.
PR 17-NOV-2000; 2000US-249213P.
PR 17-NOV-2000; 2000US-249214P.
PR 17-NOV-2000; 2000US-249215P.
PR 17-NOV-2000; 2000US-249216P.
PR 17-NOV-2000; 2000US-249217P.
PR 17-NOV-2000; 2000US-249218P.
PR 17-NOV-2000; 2000US-249244P.
PR 17-NOV-2000; 2000US-249245P.
PR 17-NOV-2000; 2000US-249264P.
PR 17-NOV-2000; 2000US-249265P.
PR 17-NOV-2000; 2000US-249297P.
PR 17-NOV-2000; 2000US-249299P.
PR 17-NOV-2000; 2000US-249300P.
PR 01-DEC-2000; 2000US-250160P.
PR 01-DEC-2000; 2000US-250391P.
PR 05-DEC-2000; 2000US-251030P.
PR 05-DEC-2000; 2000US-251988P.
PR 05-DEC-2000; 2000US-256719P.
PR 06-DEC-2000; 2000US-251479P.
PR 08-DEC-2000; 2000US-251856P.
PR 08-DEC-2000; 2000US-251868P.
PR 08-DEC-2000; 2000US-251869P.
PR 08-DEC-2000; 2000US-251989P.
PR 11-DEC-2000; 2000US-254097P.

PR 05-JAN-2001; 2001US-259678P.
PA (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
PI MPI: 2001-476161/51.
DR P-PSDB; ABB10249.
XX
XX Isolated nucleic acid molecule encoding an inflammation-associated
PT polypeptide is used in preventing, treating or ameliorating a medical
PT condition
PS Claim 1; SEQ ID NO: 137; 859pp + Sequence listing; English.
XX
XX The present invention provides human cDNAs, proteins and related genomic
CC DNAs. These can be used in the treatment of neural, immune system,
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC renal and proliferative disorders and inflammation. The present sequence
CC is a cDNA of the invention.
XX
SQ Sequence 530 BP; 158 A; 105 C; 95 G; 171 T; 1 other:

Query Match 14.7%; Score 528.6; DB 22; Length 530;

Best Local Similarity 99.8%; Pred. No. 1.1e-113;

Matches 528; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1744 gccacagattatcaatggtgcaattttctgtgataattggcgcgattatcatcata 1803
DB 2 gccacagattatcaatggtgcaattttctgtgataattggcgcgattatcatcata 61
QY 1804 gtttttccatggaagcagtggtttatagtgatcatcaaatgycataacagcaactgaa 1863
DB 62 gtttttccatggaagcagtggtttatagtgatcatcaaatgycataacagcaactgaa 121
QY 1864 atacgaatcaaatgtaaaaaagagatgaccccttgccaaagcttttcttatagtatt 1923
DB 122 atacgaatcaaatgtaaaaaagagatgaccccttgccaaagcttttcttatagtatt 181
QY 1924 actgatgcattatgctggaatccattttgtatgtaaatcttccactgcttcagta 1983
DB 182 actgatgcattatgctggaatccattttgtatgtaaatcttccactgcttcagta 241
QY 1984 gaaataccaggtacccaataaccttgggtagtgatttatttccaccataacagtgc 2043
DB 242 gaaataccaggtacccaataaccttgggtagtgatttatttccaccataacagtgc 301
QY 2044 ttgaaccgaattctctatctctgacacacagaagcatttaagaatgattatcggtt 2103
DB 302 ttgaaccgaattctctatctctgacacacagaagcatttaagaatgattatcggtt 361
QY 2104 tggataactacagacaaaataatctatgagacagcaaggttcagaaaacatatgctcca 2163
DB 362 tggataactacagacaaaataatctatgagacagcagwggttcagaaaacatatgctcca 421
QY 2164 tcatlcatctgggtggaatgtyggccactgycaggaagatgycacactgaatgaagccg 2223
DB 422 tcatlcatctgggtggaatgtyggccactgycaggaagatgycacactgaatgaagccg 481
QY 2224 gacctttcacatacccttggaatgtaatgtaactgatttccatacaaga 2272
DB 482 gacctttcacatacccttggaatgtaatgtaactgatttccatacaaga 530

RESULT 11

ID AAS28950 standard; cDNA; 530 BP.

XX AAS28950;

XX 21-NOV-2001 (first entry)

XX

DE cDNA encoding for human uterine motility-association polypeptide #15.
XX
KW Human; uterine motility-association disorder; uterus; pregnancy;
KM labour; menstrual cycle; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN W0200155201-A1.
PD
XX 02-AUG-2001.
PF 17-JAN-2001; 2001MO-US01317.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 06-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.

PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Barash SC, Ruben SM;
DR WPI: 2001-488777/53.
DR P-PSDB: AAU18108.
XX
XX Isolated polypeptide and nucleic acid molecules for treating,
PT preventing and/or prognosing disorders related to uterine motility
PT e.g. disorders associated with pregnancy and the menstrual cycle -
XX
PS Claim 4; SEQ ID No 25; 524pp; English.
XX
CC The present invention relates to the isolation of novel human
CC uterine motility-association polypeptides (AAU18094-AAU18152),
CC and cDNA and genomic sequences encoding for these polypeptides.
CC The sequences of the invention are useful in the diagnosis,
CC treatment, prevention and/or prognosis of diseases associated
CC with uterine motility such as pregnancy and labour, and menstrual
CC disorders. The polynucleotide sequences of the invention are also
CC useful in gene therapy. AAS28936-AAS28994 represent cDNA sequences
CC encoding for novel human uterine motility-association polypeptides.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 530 BP; 158 A; 105 C; 95 G; 171 T; 1 other:
SQ

Query Match 14.7%; Score 528.6; DB 22; Length 530;
Best Local Similarity 99.8%; Pred. NO. 1.1e-113;
Matches 528; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
0;

QY 1744 gccacagattatcagtggaattttctgtgataatttggcgcgcatatcatcata 1803
DB 2 gccacagattatcagtggaattttctgtgataatttggcgcgcatatcatcata 61
QY 1804 gtttttctcatggaagcagtttttatagtgatcatcaaatggtccataacagcaactgaa 1863
DB 62 gtttttctcatggaagcagtttttatagtgatcatcaaatggtccataacagcaactgaa 121
QY 1864 atacggaatcaagttaaaaaagagatgatcctgtgcaaacggttttttcttataagattt 1923
DB 122 atacggaatcaagttaaaaaagagatgatcctgtgcaaacggttttttcttataagattt 181
QY 1924 actgatgcatatgctggaataccattttgtatggaatttcttcaactgctcaggtta 1983
DB 182 actgatgcatatgctggaataccattttgtatggaatttcttcaactgctcaggtta 241
QY 1984 gaaatcccggtaccataactcttggtagtgatatttttttgcgccataacagtgct 2043
DB 242 gaaatcccggtaccataactcttggtagtgatatttttttgcgccataacagtgct 301
QY 2044 ttgaacccaattctctatactctgaccacaagccatttaagaatgattatcggttt 2103
DB 302 ttgaacccaattctctatactctgaccacaagccatttaagaatgattatcggttt 361
QY 2104 tggataactacagacaagaagaatctatgagacagcaaggttcagaaacataatgctcca 2163
DB 362 tggataactacagacaagaagaatctatgagacagcaaggttcagaaacataatgctcca 421
QY 2164 tcatcatctgggtggaatagtgccacgagagagatgacacactgagttatgaagcgg 2223
DB 422 tcatcatctgggtggaatagtgccacgagagagatgacacactgagttatgaagcgg 481
QY 2224 gacctttcacatcacccctgtgaaatgtaactatcttcaatcaacaga 2272
DB 482 gacctttcacatcacccctgtgaaatgtaactatcttcaatcaacaga 530

RESULT 12
AAS29573
ID AAS29573 standard; cDNA; 530 BP.
XX
XX AAS29573;
AC
AC AAS29573;
DT
DT 21-NOV-2001 (first entry)
XX
XX Human endocrine polypeptide encoding cDNA SEQ ID No 73.
DE
XX
XX Endocrine protein; human; mouse; rabbit; goat; horse; food additive;
KW cat; dog; chicken; sheep; immunosuppressive; antiarthritic; vasotropic;
KW antihemetic; antiproliferative; cytostatic; cardiant; neuroprotective;
KW cerebroprotective; nootropic; antibacterial; vitrucide; fungicide; cancer;
KW ophthalmological; vulnereary; gene therapy; autoimmune disease; neoplasm;
KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ss;
KW cerebrovascular disorder; nervous system disorder; bacterial infection;
KW fungal infection; viral infection; ocular disorder; endocrine disorder;
KW gastrointestinal disorder; renal disorder; respiratory disorder;
KW wound healing; skin aging; organ transplantation; food preservative;
KW tissue regeneration; anti-infertility.
XX
XX Homo sapiens.
OS
OS
PN
PN W0200155364-A2.
XX
XX 02-AUG-2001.
PD
XX
XX 17-JAN-2001; 2001MO-US01308.
PE
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189674.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228524.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.

```
PR 06-SEP-2000: 2000US-0230437.
PR 06-SEP-2000: 2000US-0230438.
PR 08-SEP-2000: 2000US-0231242.
PR 08-SEP-2000: 2000US-0231243.
PR 08-SEP-2000: 2000US-0231244.
PR 08-SEP-2000: 2000US-0231413.
PR 08-SEP-2000: 2000US-0231414.
PR 08-SEP-2000: 2000US-0232080.
PR 08-SEP-2000: 2000US-0232081.
PR 12-SEP-2000: 2000US-0231968.
PR 14-SEP-2000: 2000US-0232397.
PR 14-SEP-2000: 2000US-0232398.
PR 14-SEP-2000: 2000US-0232399.
PR 14-SEP-2000: 2000US-0232400.
PR 14-SEP-2000: 2000US-0232401.
PR 14-SEP-2000: 2000US-0233063.
PR 14-SEP-2000: 2000US-0233064.
PR 14-SEP-2000: 2000US-0233065.
PR 21-SEP-2000: 2000US-0234223.
PR 21-SEP-2000: 2000US-0234274.
PR 25-SEP-2000: 2000US-0234997.
PR 25-SEP-2000: 2000US-0234998.
PR 26-SEP-2000: 2000US-0235484.
PR 27-SEP-2000: 2000US-0235834.
PR 27-SEP-2000: 2000US-0235836.
PR 29-SEP-2000: 2000US-0236327.
PR 29-SEP-2000: 2000US-0236367.
PR 29-SEP-2000: 2000US-0236368.
PR 29-SEP-2000: 2000US-0236369.
PR 29-SEP-2000: 2000US-0236370.
PR 02-OCT-2000: 2000US-0236802.
PR 02-OCT-2000: 2000US-0237037.
PR 02-OCT-2000: 2000US-0237038.
PR 02-OCT-2000: 2000US-0237039.
PR 02-OCT-2000: 2000US-0237040.
PR 13-OCT-2000: 2000US-0239935.
PR 13-OCT-2000: 2000US-0239937.
PR 20-OCT-2000: 2000US-0240960.
PR 20-OCT-2000: 2000US-0241221.
PR 20-OCT-2000: 2000US-0241785.
PR 20-OCT-2000: 2000US-0241786.
PR 20-OCT-2000: 2000US-0241787.
PR 20-OCT-2000: 2000US-0241808.
PR 20-OCT-2000: 2000US-0241809.
PR 20-OCT-2000: 2000US-0241826.
PR 01-NOV-2000: 2000US-0244617.
PR 08-NOV-2000: 2000US-0246474.
PR 08-NOV-2000: 2000US-0246475.
PR 08-NOV-2000: 2000US-0246476.
PR 08-NOV-2000: 2000US-0246477.
PR 08-NOV-2000: 2000US-0246478.
PR 08-NOV-2000: 2000US-0246523.
PR 08-NOV-2000: 2000US-0246524.
PR 08-NOV-2000: 2000US-0246525.
PR 08-NOV-2000: 2000US-0246526.
PR 08-NOV-2000: 2000US-0246527.
PR 08-NOV-2000: 2000US-0246528.
PR 08-NOV-2000: 2000US-0246532.
PR 08-NOV-2000: 2000US-0246609.
PR 08-NOV-2000: 2000US-0246610.
PR 08-NOV-2000: 2000US-0246611.
PR 08-NOV-2000: 2000US-0246613.
PR 17-NOV-2000: 2000US-0249207.
PR 17-NOV-2000: 2000US-0249208.
PR 17-NOV-2000: 2000US-0249209.
PR 17-NOV-2000: 2000US-0249210.
PR 17-NOV-2000: 2000US-0249211.
PR 17-NOV-2000: 2000US-0249212.
PR 17-NOV-2000: 2000US-0249213.
PR 17-NOV-2000: 2000US-0249214.
PR 17-NOV-2000: 2000US-0249215.
PR 17-NOV-2000: 2000US-0249216.
PR 17-NOV-2000: 2000US-0249217.
PR 17-NOV-2000: 2000US-0249217.

PR 17-NOV-2000: 2000US-0249218.
PR 17-NOV-2000: 2000US-0249244.
PR 17-NOV-2000: 2000US-0249245.
PR 17-NOV-2000: 2000US-0249264.
PR 17-NOV-2000: 2000US-0249265.
PR 17-NOV-2000: 2000US-0249265.
PR 17-NOV-2000: 2000US-0249299.
PR 17-NOV-2000: 2000US-0249300.
PR 01-DEC-2000: 2000US-0250160.
PR 01-DEC-2000: 2000US-0250391.
PR 05-DEC-2000: 2000US-0251030.
PR 05-DEC-2000: 2000US-0251988.
PR 05-DEC-2000: 2000US-0256719.
PR 06-DEC-2000: 2000US-0251479.
PR 08-DEC-2000: 2000US-0251856.
PR 08-DEC-2000: 2000US-0251868.
PR 08-DEC-2000: 2000US-0251869.
PR 08-DEC-2000: 2000US-0251869.
PR 08-DEC-2000: 2000US-0251989.
PR 11-DEC-2000: 2000US-0254097.
PR 05-JAN-2001: 2001US-0259678.

XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PA
XX PI Rosen CA, Barash SC, Ruben SM;
XX PI WPI: 2001-451936/48.
XX PI P-PSDB; AAU18344.
XX DR
XX DR
XX PT Isolated polypeptide for treating, preventing and/or prognosing
XX PT disorders of the endocrine system such as reproductive disorders,
XX PT endocrine cancers and also for testing and detection e.g. diagnosis -
XX PS Claim 1; SEQ ID NO 73; 604bp; English.
XX PS
XX CC Sequences AAS29511-AAS29736 represent cDNA molecules, which encode the
XX CC endocrine polypeptides of the invention. Endocrine polypeptides and their
XX CC associated polynucleotides of the invention are useful in the diagnosis,
XX CC treatment and prevention of various types of disorders in e.g. humans,
XX CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. A
XX CC pathological condition can be determined by determining the presence or
XX CC absence of a mutation in an endocrine polynucleotide. The treatable
XX CC disorders include autoimmune diseases such as rheumatoid arthritis,
XX CC hyperproliferative disorders such as neoplasms of the breast or liver,
XX CC cardiovascular disorders such as cardiac arrest, cerebrovascular
XX CC disorders such as cerebral ischaemia, nervous system disorders such as
XX CC Alzheimer's disease, infections caused by bacteria, viruses and fungi,
XX CC ocular disorders such as corneal infection, endocrine disorders such as
XX CC premature labour and infertility, gastrointestinal disorders such as
XX CC Crohn's disease, renal disorders such as glomerulonephritis and
XX CC respiratory disorders such as asthma. The polypeptides can also be used
XX CC to aid wound healing, to prevent skin aging due to sunburn, to maintain
XX CC organs before transplantation, to regenerate tissues and in chemotaxis.
XX CC The polypeptides can also be used as a food additive or preservative to
XX CC increase or decrease storage capabilities.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO

Query Match 14.7%; Score 528.6; DB 22; Length 530;
Best Local Similarity 99.8%; Pred. No. 1.1e-113;
Matches 528; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1744 gccacgattatcagtggaacatttctctgtatattggtggcgcattatcata 1803
DB 2 gccacgattatcagtggaacatttctctgtatattggtggcgcattatcata 61
QY 1804 gtttttccatgaagaattttatagttttatcaaaatgccaataagcaacgaa 1863
DB 62 gtttttccatgaagaattttatagttttatcaaaatgccaataagcaacgaa 121
QY 1864 atacgaatacaagtttaaaagagatgatccttgccaacgtttttttatagttt 1923
DB 122 atacgaatacaagtttaaaagagatgatccttgccaacgtttttttatagttt 181
```

OY 1924 actgatgcattatgctgtagtaccattttgtgtagtgaattcttctcaactgctcaagta 1983
|||||
Db 182 actgatgcatltagctgtagataccattttgtgtagtgaattcttctcaactgctcaagta 241
OY 1984 gaaataccaggtacataaactcttggtgtagtattttattctgcacataaactgct 2043
|||||
Db 242 gaaataccaggtacataaactcttggtgtagtattttattctgcacataaactgct 301
OY 2044 tgaacccaattctctatctctgaccacacagccatttaaaagtatgctgttt 2103
|||||
Db 302 tgaacccaattctctatctctgaccacacagccatttaaaagtatgctgttt 361
OY 2104 tggataactacagacaagaataatctatgacagcaaggtcagaanaacatgctcca 2163
|||||
Db 362 tggataactacagacaagaataatctatgacagcaaggtcagaanaacatgctcca 421
OY 2164 tcatcatctggtgtagtgaatgtagcactgacagagatgacacctgagttaagccg 2223
|||||
Db 422 tcatcatctggtgtagtgaatgtagcactgacagagatgacacctgagttaagccg 481
OY 2224 gacctttcacatacccttgtagaagtgcactgatttctcaatcaacga 2272
|||||
Db 482 gacctttcacatacccttgtagaagtgcactgatttctcaatcaacga 530

RESULT 13

AAS30187
ID AAS30187 standard; cDNA; 530 BP.

AC AAS30187;
XX

DT 21-NOV-2001 (first entry)
XX

DE DNA encoding renal and cardiovascular-associated protein, Seq ID 33.
XX

XX Human; antiinflammatory; neuroprotective; immunomodulator; vulnereary;
cardiovascular; cytostatic; nephrotoxic; antianaemic; nephritis;
KW Immunosuppressive; kidney disorder; renal failure; hypertension;
KW cardiovascular disorder; myocardial infarction; blood disorder; anaemia;
KW blood coagulation disorder; electrolyte imbalance disorder; cancer;
KW hyponatraemia; hyperkalaemia; neoplastic disorder; nephroma;
KW autoimmune disease; inflammatory disease; reproductive system disorder;
KW endocrine disorder; neural activity; neurological disorder;
KW wound healing; respiratory disorder; ss.

OS Homo sapiens.
XX

PN WO200155328-A2.
XX

PD 02-AUG-2001.
XX

PF 17-JAN-2001; 2001WO-US01359.
XX

XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-MAR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216680.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0227824.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234597.
PR 25-SEP-2000; 2000US-0234598.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239335.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240360.
PR 20-OCT-2000; 2000US-0241121.
PR 20-OCT-2000; 2000US-0241185.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0244675.
PR 08-NOV-2000; 2000US-0244676.
PR 08-NOV-2000; 2000US-0246477.

PR	08-NOV-2000	2000US-0246178.
PR	08-NOV-2000	2000US-0246523.
PR	08-NOV-2000	2000US-0246524.
PR	08-NOV-2000	2000US-0246525.
PR	08-NOV-2000	2000US-0246526.
PR	08-NOV-2000	2000US-0246527.
PR	08-NOV-2000	2000US-0246528.
PR	08-NOV-2000	2000US-0246532.
PR	08-NOV-2000	2000US-0246609.
PR	08-NOV-2000	2000US-0246610.
PR	08-NOV-2000	2000US-0246611.
PR	08-NOV-2000	2000US-0246613.
PR	17-NOV-2000	2000US-0249207.
PR	17-NOV-2000	2000US-0249208.
PR	17-NOV-2000	2000US-0249209.
PR	17-NOV-2000	2000US-0249210.
PR	17-NOV-2000	2000US-0249211.
PR	17-NOV-2000	2000US-0249212.
PR	17-NOV-2000	2000US-0249213.
PR	17-NOV-2000	2000US-0249214.
PR	17-NOV-2000	2000US-0249215.
PR	17-NOV-2000	2000US-0249216.
PR	17-NOV-2000	2000US-0249217.
PR	17-NOV-2000	2000US-0249218.
PR	17-NOV-2000	2000US-0249244.
PR	17-NOV-2000	2000US-0249245.
PR	17-NOV-2000	2000US-0249264.
PR	17-NOV-2000	2000US-0249265.
PR	17-NOV-2000	2000US-0249297.
PR	17-NOV-2000	2000US-0249299.
PR	17-NOV-2000	2000US-0249300.
PR	01-DEC-2000	2000US-0250160.
PR	01-DEC-2000	2000US-0250391.
PR	05-DEC-2000	2000US-0251030.
PR	05-DEC-2000	2000US-0251988.
PR	05-DEC-2000	2000US-0256719.
PR	06-DEC-2000	2000US-0251479.
PR	08-DEC-2000	2000US-0251856.
PR	08-DEC-2000	2000US-0251868.
PR	08-DEC-2000	2000US-0251869.
PR	08-DEC-2000	2000US-0251989.
PR	08-DEC-2000	2000US-0251990.
PR	11-DEC-2000	2000US-0254097.
PR	05-JAN-2001	2001US-0259678.
XX		
PA	(HUMA-)	HUMAN GENOME SCI INC.
PI	Rosen CA,	Barash SC, Ruben SM;
PI	WPI;	2001-488787/53.
XX	P-PSDB;	AAU18666.
DR		
XX		
PT	New polynucleotides and polypeptides,	useful for diagnosing, treating,
PT	preventing or prognosing e.g. kidney,	cardiovascular, blood,
PT	electrolyte imbalance or neoplastic disorders,	autoimmune diseases,
PT	cancers	
XX		
PS	Claim 1;	SEQ ID No 33; 506bp; English.
XX		
CC	The invention relates to novel nucleic acids and polypeptides useful for	
CC	diagnosing, treating, preventing and/or prognosing disorders related to	
CC	the various polypeptides. The polynucleotides are especially useful in the	
CC	diagnosis, prognosis, prevention and/or treatment of diseases which	
CC	include kidney disorders (e.g. renal failure or nephritis),	
CC	cardiovascular disorders (e.g. hypertension or myocardial infarction),	
CC	blood disorders (e.g. anaemia or blood coagulation disorders),	
CC	electrolyte imbalance disorders (e.g. hyponatraemia or hyperkalaemia),	
CC	neoplastic disorders (e.g. nephroma or renal cell cancer), autoimmune	
CC	diseases, cancers, inflammatory diseases, reproductive system	
CC	disorders, endocrine disorders, neural activity and neurological	
CC	disorders, wound healing and respiratory disorders. AAS30165-AAS30251	
CC	represent the novel human renal and cardiovascular-associated nucleic	
CC	acid sequences of the invention. Note: The sequence data for this patent	
CC		

CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at:
CC [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).
XX
SQ Sequence 530 BP; 158 A; 105 C; 95 G; 171 T; 1 other;

Query Match	14.7%	Score 528.6;	DB 22;	Length 530;
Best Local Similarity	99.8%;	Pred. No. 1.1e-113;		
Matches 528; Conservative	1;	Mismatches 0;	Indels 0;	Gaps 0;

QY 1744 gccacgattatcatcagtcgcaattttctcttgtaataattggccgcatattatcatcata 1803

QY 1804 gttttttcctatgtaagcagtgttttatagtttcatcaaatgtccataacagcaactgaa 1863
Db 62 gttttttcctatgtaagcagtgttttatagtttcatcaaatgtccataacagcaactgaa 121

QY 1864 ataaaggaatccaagttaaaaaagagatgatccttgcacaagtttttcttatagtatt 1923
|||||
Db 122 ataaaggaatccaagttaaaaaagagatgatccttgcacaagtttttcttatagtatt 181

Qy 1924 acgcatgcatatctctgatacccatlctgttagtgaatctcttcaactgcttcagta 1983
 |||||
 Db 182 acgatgcatatgctgatacccatlctgttagtgaatctcttcaactgcttcagta 241

QY 1984 gaataaccaggtacacataactctctggtagtgattttatctgcacattaacagtcgct 2043
|||||
Db 242 gaatatcccggtacacataactctctggtagtgattttatctgcacattaacagtcgct 301

2074 ccgaaccccaatctctatactctgacacaaagaccattaaagaaatgattcactggttt 361

Db 362 tggcataactacagacaagaataatctatgtgacagcagwagtcagaaacatctgtcca 421

Db 422 tcaatcatctggtggaatgtgcccactgcagagatgccactgagttaatgaagcgc 481

Db 482 gaccttccataccocctggaargtactcagatttcacatcaagca 530

AAS34845
 ID AAS34845 standard; cDNA; 530 BP.
 XX
 AC AAS34845.

XX	04-DEC-2001	(first entry)
DT		
XX		
DE	CDNA encoding novel human neoplastic disease associated polypeptide #77	

KW Human; neoplastic disease associated polypeptide; cancer; gene therapy
KW hyperproliferative disorder; neural disorder; immune system disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;

KW neuroprotective; cytostatic; anti inflammatory; vasotropic; ss.
XX
OS Homo sapiens.
XX

PN WOZ00133103-AL.
XX
PD 02-AUG-2001.
XX

XX
PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235634.
PR 27-SEP-2000; 2000US-0235636.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.

PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241285.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-465558/50.
DR P-PSDB; AUG21646.
XX
PT Novel polypeptides and polynucleotides useful as diagnostic reagents to
PT diagnose diseases or disorders associated with aberrant expression or
PT activity of polypeptides, and for treating cancers, rheumatoid

PT arthritis -
XX
PS Claim 4; SEQ ID NO 89; 687pp; English.
XX
CC The present invention relates to the isolation of novel human neoplastic
CC disease associated polypeptides (AAU21568-AAU21551), and cDNA and DNA
CC sequences encoding for these polypeptides. The sequences of the
CC invention are useful in the diagnosis, treatment, prevention and/or
CC prognosis of disorders involving neoplastic disease such as
CC hyperproliferative disorders (e.g. leukemia, bone cancer, bladder
CC cancer, brain stem glioma, adult liver cancer, childhood cerebellar
CC astrocytoma, or Hodgkin's lymphoma). The sequences of the invention may
CC also be useful for treating other disorders such as neural disorders,
CC immune system disorders, muscular disorders, reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC and renal disorders. The polynucleotide sequences of the invention are
CC also useful in gene therapy. AAS34767-AAS35050 represent cDNA sequences
CC encoding for the novel human neoplastic disease associated polypeptides
CC of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pubd/published_pct_sequences.
XX
SQ Sequence 530 BP; 158 A; 105 C; 95 G; 171 T; 1 other;

Query Match 14.7%; Score 528.6; DB 22; Length 530;
Best Local Similarity 99.8%; Pred. No. 1.1e-113;
Matches 528; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1744 gccacagattatcagtcggaattttctgtgattattgtggcgcattatcatcata 1803
DB 2 gccacagattatcagtcggaattttctgtgattattgtggcgcattatcatcata 61

QY 1804 gtttttccctatggaagcagctgtttatagtgatcatcaaaagtcgcatcaagcaactgaa 1863
DB 62 gtttttccctatggaagcagctgtttatagtgatcatcaaaagtcgcatcaagcaactgaa 121

QY 1864 atcgcgaatcaagttaaaaaagagatgatccttcgccaaggttttttctttagtattt 1923
DB 122 atcgcgaatcaagttaaaaaagagatgatccttcgccaaggttttttctttagtattt 181

QY 1924 actgagcatltagctcgatgataccattttgtatgtgaatttcttcaactgctcagtt 1983
DB 182 actgagcatltagctcgatgataccattttgtatgtgaatttcttcaactgctcagtt 241

QY 1984 gaaatccaggtacatcaactcttggtgtagtattttatcttgcacattaaacagtct 2043
DB 242 gaaatccaggtacatcaactcttggtgtagtattttatcttgcacattaaacagtct 301

QY 2044 ttgaaccgaattctctatctctgaccacagaacatttaaaaaatgattatcggtt 2103
DB 302 ttgaaccgaattctctatctctgaccacagaacatttaaaaaatgattatcggtt 361

QY 2104 tggatataactacagacaagaataatctatgacagcaaaagtcagaaaacatctgcaca 2163
DB 362 tggatataactacagacaagaataatctatgacagcaaaagtcagaaaacatctgcaca 421

QY 2164 tcaatcatctgggtggaatgtggtgcaactgacagagatgacacactgaattgaagcgcg 2223
DB 422 tcaatcatctgggtggaatgtggtgcaactgacagagatgacacactgaattgaagcgcg 481

QY 2224 gaccttttcacataccctctggaatgtcactgatttctcaatcaacga 2272
DB 482 gaccttttcacataccctctggaatgtcactgatttctcaatcaacga 530

DT 08-JAN-2002 (first entry)
XX
DE Human musculoskeletal system related polynucleotide SEQ ID NO 917.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
KW vulnerrary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein;
KW musculoskeletal system; ss.
XX
OS Homo sapiens.
XX
PN MO200155367-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01338.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216680.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225477.
PR 14-AUG-2000; 2000US-0225478.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.

Db 263 CTCCTTGCACTGCCAGCAGGAGATACTACTTTTAACTAGTTGTTCTTCCTAGTCT 204
QY 3415 ctacgttattagnatttttggcttcataatgtaaacctttaagcaggagaagaaatg 3474
Db 203 CTACGTTATTAGAAATTTTGGCTTTCATTAAGTGAAACCTTTAAGCAGGAGAGAGAAAAATG 144
QY 3475 ttctcagatagttccaatacncacaaatgttgcaacacacaaatcactcgaatcnaac 3534
Db 143 TTTTCAGATAGTTTCAAAATACACCAAAAATGTTGAAACACAAAAATACTGGAATCAAAC 84
QY 3535 cataatgcccttaltgaatatatagttgctatagntttgcctgaaaa 3581
Db 83 CATAAATGCACCTTATTGAATATATAGTTGTATAGATTGTTCTGAAAA 37

Search completed: September 5, 2002, 08:24:32
Job time: 16102 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2002, 02:59:25 : Search time 112.92 Seconds
(without alignments)
7796.232 Million cell updates/sec

Title: US-09-647-067-7

Perfect score: 3584

Sequence: 1 cgccttgcgaactgactgaaga.....tagntttgtctgaaaccc 3584

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 38353 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA: *
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata/2/ina/Backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59.8	1.7	3573	4	Sequence 4, Appl
2	59.8	1.7	6471	4	Sequence 1, Appl
3	58.8	1.6	2290	4	Sequence 4, Appl
4	54.8	1.5	4417	4	Sequence 57, Appl
5	50.8	1.4	636	4	Sequence 1137, Ap
6	50.6	1.4	2179	4	Sequence 1, Appl
7	50.6	1.4	2179	2	Sequence 1, Appl
8	50.6	1.4	2179	3	Sequence 1, Appl
9	49.4	1.4	837	4	Sequence 288, App
10	47.6	1.3	2022	4	Sequence 9, Appl
11	47.6	1.3	2987	4	Sequence 1, Appl
12	47.6	1.3	3979	4	Sequence 1, Appl
13	47.6	1.3	3979	4	Sequence 2, Appl
14	47.6	1.3	4213	4	Sequence 7, Appl
15	47.6	1.3	7218	4	Sequence 14, Appl
16	46.8	1.3	3710	4	Sequence 62, Appl
17	46.8	1.3	8378	5	Sequence 1, Appl
18	46.4	1.3	426	1	Sequence 35, Appl
19	46.4	1.3	567	1	Sequence 38, Appl
20	46.4	1.3	711	1	Sequence 41, Appl
21	46.4	1.3	849	1	Sequence 44, Appl
22	46.4	1.3	924	1	Sequence 47, Appl
23	46.4	1.3	1002	1	Sequence 26, Appl
24	46.4	1.3	1026	1	Sequence 1, Appl
25	46.4	1.3	1026	1	Sequence 1, Appl
26	46.4	1.3	1026	5	Sequence 1, Appl
27	46.4	1.3	1593	6	Sequence 1, Appl

28	44.2	1.2	4203	2	US-08-866-757-1	Sequence 1, Appl
29	44.2	1.2	4203	4	US-09-153-593-1	Sequence 1, Appl
30	44	1.2	2317	3	US-08-749-522-5	Sequence 5, Appl
31	44	1.2	3974	4	US-08-467-504-3	Sequence 3, Appl
32	43.8	1.2	2296	3	US-09-188-930-228	Sequence 28, App
33	43.6	1.2	615	4	US-08-998-416-186	Sequence 16, App
34	43.6	1.2	19124	2	US-08-487-826B-13	Sequence 13, Appl
35	43.4	1.2	2513	4	US-09-228-986-13	Sequence 1, Appl
36	43.4	1.2	51952	3	US-08-947-823-1	Sequence 1, Appl
37	43	1.2	4758	3	US-09-191-647-1	Sequence 1, Appl
38	43	1.2	4758	4	US-09-540-245A-1	Sequence 1, Appl
39	43	1.2	4758	4	US-09-540-153-1	Sequence 1, Appl
40	42.8	1.2	2168	3	US-08-749-522-6	Sequence 6, Appl
41	42	1.2	1677	2	US-08-684-101-1	Sequence 1, Appl
42	42	1.2	1677	4	US-09-205-814-1	Sequence 1, Appl
43	42	1.2	5176	4	US-09-182-024A-1	Sequence 1, Appl
44	41.6	1.2	3925	2	US-09-047-026A-3	Sequence 3, Appl
45	41.6	1.2	6854	1	US-08-460-036-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-09-353-585-4
Sequence 4, Application US/09353585
Patent No. 6287865
GENERAL INFORMATION:
APPLICANT: Dixon, Mark S
Jones, Jonathan DG
TITLE OF INVENTION: Plant pathogen resistance genes and uses
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESS: 8th Floor, 1100 No. 6287865th Glebe Road
CITY: Arlington
STATE: Virginia
COUNTRY: United States of America
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/353,585
FILING DATE: 15-Jul-1999
CLASSIFICATION: C12N 15/29, 15/82, A01H 5/00, A01N 65/00, C12O 1/68
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/930,277
FILING DATE: 27-OCT-1997
APPLICATION NUMBER: PCT/GB96/00785
FILING DATE: 01-APR-1996
APPLICATION NUMBER: GB 9506558.5
ATTORNEY/AGENT INFORMATION:
NAME: Ms Mary J Wilson
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 620-69
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3573 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO

```

; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Tomato
; STRAIN: CF2
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-353-585-4

```

```

Query Match      1.7%; Score 59.8; DB 4; Length 3573;
Best Local Similarity 45.1%; Pred. No. 2.1e-05;
Matches 263; Conservative 0; Mismatches 317; Indels 3; Gaps 1;

```

```

QY 424 aacttaagaagaagctccctcgtatgcttcaagaatcatcagatcctcagaagctg 483
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 521 AACCTTCTTAAGTGGTCCATTCCTCCCTCAGTGGGGAATCGAACACTTCTTTTGG 580
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 484 gaccggcaacaataagattacatccatcccatctatgtcttcaaggagctgatgc 543
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 501 TATCTTTACAAATACACCTTCTGGCTCTATTCCTGAAGAAATAGTTACCTAAGATCT 640
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 544 ctactaaactgtatcctcagatcatacagaataaccctccggaagcgagtgcttga 603
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 641 CTTACTGAGCTAGANTTGTAGATATATGCTCTATATGCGTCTATTCCTTCATTTGGG 700
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 604 gatcttcacagactagaatgagctgataattgaagaatacaccctcagtcgaattccca 663
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 701 AATATGAAACAATCTCTTTTGTGTTTCTTATGGAATACGCTTTCGCTCTATTCCT 760
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 664 ccaacattttagactaaatctcttcttcttcttcttcttcttcttcttcttcttcc 723
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 761 GAGAAATATATCTTACCTAGATCTTCTTACTTACCTAGATTTGAGTGAAGATGCTTAT 820
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 724 cgttaccgtataaacctctctgtcaacacatgccaagactacatctgctgagacttga 783
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 821 GGCT---CTATTCCTGCTTCAATTTGGGAATTTGAACAACCTGTCTTTTGTCTTTAT 877
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 784 ggcacacatcatcaatattagaattgacttatttctcagtaattcaactgtt 843
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 878 GGAATACGCTTTCTGGCTCTATTCCTGAAGAAATAGCTTACCTAAGATCTTATTC 937
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 844 ttatgtatgagaagaacaaataatcaactcaactaaatcaacttgcacccctccag 903
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 938 CTAGTTTGAATGAGAAATGCTTCTTATATGCTCTATTCCTGCTTCTTTGGGAATTCGAA 997
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 904 aaactgaatgaatgattgaagaatgaatgaatgaatgaatgaatgaatgaatgaatga 963
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 998 AACTTGTCTAGTGAATCTTGTAAATATCAGCTTCTGCTCTATTCCTGCTTCATTTG 1057
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 964 aagacccgaagagcgctgcacaaattgaatttccctatac 1006
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1058 GGGAAATCGAACACTTGTCTATGCTTATCTTAAACAATAAC 1100
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 2
US-09-353-585-1
; Sequence 1, Application US/09353585
; Patent No. 6287/865
; GENERAL INFORMATION:
; APPLICANT: Dixon, Mark S
; Jones, David A
; Jones, Jonathan DG
; TITLE OF INVENTION: Plant pathogen resistance genes and uses
; thereof
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye PC
; STREET: 8th Floor, 1100 No. 6287865th Giebe Road
; CITY: Arlington
; STATE: Virginia
; COUNTRY: United States of America
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/353,585
; FILING DATE: 15-Jul-1999
; CLASSIFICATION: C12N 15/29, 15/82, A01H 5/00, A01N 65/00, C12Q
; 1/68
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/930,277
; FILING DATE: 27-Oct-1997
; APPLICATION NUMBER: PCT/GB96/00785
; FILING DATE: 01-Apr-1996
; APPLICATION NUMBER: GB 9506658.5
; FILING DATE: 31-Mar-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Ms. Mary J. Wilson
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 620-69
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ. ID NO. 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6471 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Tomato
; STRAIN: CF2
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1754..5012
; NAME/KEY: sig_peptide
; LOCATION: 1677..1753
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-353-585-1

```

```

Query Match      1.7%; Score 59.8; DB 4; Length 6471;
Best Local Similarity 45.1%; Pred. No. 2.7e-05;
Matches 263; Conservative 0; Mismatches 317; Indels 3; Gaps 1;
QY 424 aacttaagaagaagctccctcgtatgcttcaagaatcatcagatcctcagaagctg 483
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2205 AACTTCTTAAGTGGTCCATTCCTCCCTCAGTGGGGAATCGAACACTTCTTTTGG 2264
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 484 gaccggcaacaataagattacatccatcccatctatgtcttcaaggagctgaatgc 543
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2265 TATCTTTACAAATACACCTTCTGGCTCTATTCCTGAAGAAATAGTTACCTAAGATCT 2324
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 544 ctactaaactgtatcctcagatcatacagaataaccctccggaagcgagtgcttga 603
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2325 CTTACTGAGCTAGANTTGTAGATATATGCTCTATATGCGTCTATTCCTTCATTTGGG 2384
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 604 gatcttcacagactagaatgagctgataattgaagaataatccctcagtcgaattccca 663
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2385 AATATGAAACAATCTCTTTTGTGTTTCTTATGGAATACGCTTTCGCTCTATTCCT 2444
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 664 ccaacattttagactaaatctcttcttcttcttcttcttcttcttcttcttctcacc 723
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2445 GAGAAATATATCTTACCTAGATCTTCTTACTTACCTAGATTTGAGTGAAGATGCTTAAAT 2504
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 724 cgttaccgtataaacctctctgtcaacacatgccaagactacatctgctgagacttga 783
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2505 GGCT---CTATTCCTGCTTCAATTTGGGAATTTGAACAACCTGTCTTTTGTCTTTAT 2561
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 784 ggcacacatcatcaatattagaattgacttatttctcagtaattcaactgtt 843
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```



```

1 TELECOMMUNICATION INFORMATION:
2 TELEPHONE: (617) 723-1300
3 TELEFAX: (617) 723-8923
4 INFORMATION FOR SEQ ID NO.: 1:
5 SEQUENCE CHARACTERISTICS:
6 LENGTH: 2179
7 TYPE: Nucleic acid
8 STRANDEDNESS: Double
9 TOPOLOGY: Linear
10
11 MOLECULE TYPE: cDNA to mRNA
12
13 ORIGINAL SOURCE:
14 ORGANISM: Homo sapiens
15 TISSUE TYPE: Testis
16
17 IMMEDIATE SOURCE:
18 LIBRARY: Igtil cDNA library, Clontech #HL1010b
19 CLONE: pHFSHR11-11, pHFSHR15-6
20
21 FEATURE:
22
23 NAME/KEY: protein coding region
24 LOCATION: 75 to 2159
25
26 US-08-482-855-1

```

Query Match	1.4%	Score 50.6	DB 3	Length 2179
Best Local Similarity	45.4%	Pred. No. 0.0033		
Matches 182, Conservative	0	Mismatches 219	Indels 0	Gaps 0

Oy	1149	ccagctacgttggatgtccaccacgcttcgcgcgttaaccacaacacgtggaatttc	1208
Db	1076	ctatgactcttatgcattgaaagtgttgacgtgacgtctcccttaagccacgatgcatTTCA	1135
Oy	1209	alcctagagaatctctcttgcgaagcatattatcaagaagatattgtctggttgcaltcgc	1268
Db	1136	cccatgtgaaagatattcatgttggggttacaacaactcctacagatccttgatattgattatcagcat	1195
Oy	1269	agttaccgtgctttggaacatttttgtcaatttgcatgcgatcccttatataaggtctgagaa	1328
Db	1196	cctggcccatTCATCGGGAACATCATCATGTGCTAGTATCTCAACTACAGGCCAATATMAAT	1255
Oy	1329	caagttgatgccaatgcgtcaatctattctctgcgtgctgcgcgtcttaattggaaata	1388
Db	1236	cacagatccccaggtTCCTTATGTGCAACCTGGCCCTTGTGTGATCTTGCAATGGAAATCA	1315
Oy	1389	tttaattcgtatgcgagagcttgcactaaagtttcgttgagatacaataacatgcgcga	1448
Db	1316	cctgctgcgcattggcatcagttgcatTCATTCACCAAGAACCAATATACAACTATGCCAT	1375
Oy	1449	gctgtgagatgagatcactcatctgtcaagctgtgtagatctcttggccattctgtccacaga	1508
Db	1376	tgacTGGCAAACTGGGGGAGCGTGTGATGCTGCTGGCTTTTCACATGCTTTGCCAGTGA	1435
Oy	1509	agatcagtttactcgttaacatcttcgcacattgaaaaat	1549
Db	1436	gcctgcagctTCATCACTCTGCACGATATCACTCTGGAAAGAT	1476

```

1  RESULT          9
2  US-08-998-416-288/C
3  : Sequence 288. Application US/08998416
4  : Patent No. 6239264
5  :
6  : GENERAL INFORMATION:
7  :
8  : APPLICANT:      Philippsen, Peter
9  : APPLICANT:      Pohlmann, Rainer
10 : APPLICANT:      Steiner, Sabine
11 : APPLICANT:      Mohr, Christine
12 : APPLICANT:      Wendland, Jürgen
13 : APPLICANT:      Knechtle, Philipp
14 : APPLICANT:      Reibschung, Corinne
15 : TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSEYII
16 : TITLE OF INVENTION: AND USES THEREOF
17 : NUMBER OF SEQUENCES: 1152
18 :
19 : CORRESPONDENCE ADDRESS:
20 :
21 : ADDRESSEE: No. 6239264artis Corporation
22 : STREET: 3054 Cornwallis Road
23 :

```

CITY: Research Triangle Park
 STATE: No. 6239264th Carolina
 COUNTRY: USA
 ZIP: 27709
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/998,416
 FILING DATE: 24-DEC-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: CG 0016/97
 FILING DATE: 31-DEC-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Meigs, J. Timothy
 REGISTRATION NUMBER: 38,241
 REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 919-541-8587
 TELEFAX: 919-541-8689
 INFORMATION FOR SEQ ID NO: 288:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 837 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 ORIGINAL SOURCE:
 ORGANISM: PAG124LRP

Query Match	1.4%	Score 49.4	DB 4	Length 837
Best Local Similarity	50.6%	Pred. No.	0.0045	
Matches 119	Conservative	0	Mismatches 116	Indels 0
				Gaps 0

[illegible]

RESULT 10
 US-07-757-342D-9
 ; Sequence 9, Application US/07757342D
 ; Patent No. 6218509
 ; GENERAL INFORMATION:
 ; APPLICANT: IGARASHI, Masao
 ; MINEGISHI, Takashi
 ; NAKAMURA, Kazuo
 ; TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
 ; CUSHMAN
 ; STREET: 130 Water Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: US


```

RESULT 12
US-09-180-439-1
; Sequence 1, Application US/09180439
; Patent No. 6225532
; GENERAL INFORMATION:
; APPLICANT: Dixon, Mark S
; APPLICANT: Hatzixanthis, Kostas
; APPLICANT: Jones, David A
; APPLICANT: Jones, Jonathan DG
; TITLE OF INVENTION: Plant pathogen resistance genes and uses thereof
; FILE REFERENCE: 620 - 53
; CURRENT APPLICATION NUMBER: US/09/180,439
; CURRENT FILING DATE: 1998-12-06
; EARLIER APPLICATION NUMBER: PCT/GB97/01249
; EARLIER FILING DATE: 1997-05-08
; EARLIER APPLICATION NUMBER: GB 9609681.3
; EARLIER FILING DATE: 1996-05-09
; EARLIER APPLICATION NUMBER: GB 9619924.5
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3979
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
US-09-180-439-1

Query Match      1.3%; Score 47.6; DB 4; Length 3979;
Best Local Similarity 48.2%; Pred. No. 0.023;
Matches 134; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

QY 444 tccgtatgtctcaagaattatcatatctctcaagaagctgagacctgcaaacataaagat 503
      ||||| || || |||| | | | | | | | | | | | | | | | | | | | | | | |
DB 1778 tccgtgttcattcgggaatcgtgacaactgtttatgtgtatcttacaataatcagct 1837
      ||||| || || |||| | | | | | | | | | | | | | | | | | | | | | | |
QY 504 tacatccatccatcatatgtcttcagaggaactgaataagccttaactatgtatctcag 563
      ||||| || || |||| | | | | | | | | | | | | | | | | | | | | | | |
DB 1838 ttctgctctatctccgaagaatagttacccgttctcttaccgaactatattggg 1897
      ||||| || || |||| | | | | | | | | | | | | | | | | | | | | | | |
QY 564 tcatacagaataactctctcgtgaagccgggtgttttgaagatcttcacagactagaatg 623
      ||||| || || |||| | | | | | | | | | | | | | | | | | | | | | | |
DB 1898 taataactctcttaatgagctctatctcgttcatattgggaatctgaaacaactgtttat 1957
      ||||| || || |||| | | | | | | | | | | | | | | | | | | | | | | |
QY 624 gctgataatgaagaatacctcagtcgaactcccccacaactttatgagactaa 683
      ||||| || || |||| | | | | | | | | | | | | | | | | | | | | | | |
DB 1958 gtgtatcttcaataatcagcttctcgtctatcttccctgaagaataagttacctgag 2017
      ||||| || || |||| | | | | | | | | | | | | | | | | | | | | | | |
QY 684 ttctctatctcttagtctcgtatgaataagctcctca 721
      ||||| || || |||| | | | | | | | | | | | | | | | | | | | | | | |
DB 2018 ttctcttactgaactatttttgggtataaactctcta 2055
      ||||| || || |||| | | | | | | | | | | | | | | | | | | | | | | |

RESULT 13
US-09-180-439-2
; Sequence 2, Application US/09180439
; Patent No. 6225532
; GENERAL INFORMATION:
; APPLICANT: Dixon, Mark S
; APPLICANT: Hatzixanthis, Kostas
; APPLICANT: Jones, David A
; APPLICANT: Jones, Jonathan DG
; TITLE OF INVENTION: Plant pathogen resistance genes and uses thereof
; FILE REFERENCE: 620 - 53
; CURRENT APPLICATION NUMBER: US/09/180,439
; CURRENT FILING DATE: 1998-12-06
; EARLIER APPLICATION NUMBER: PCT/GB97/01249
; EARLIER FILING DATE: 1997-05-08
; EARLIER APPLICATION NUMBER: GB 9609681.3
; EARLIER FILING DATE: 1996-05-09
; EARLIER APPLICATION NUMBER: GB 9619924.5
; EARLIER FILING DATE: 1996-09-24

```

```

; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2
; LENGTH: 3979
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
US-09-180-439-2

Query Match      1.3%; Score 47.6; DB 4; Length 3979;
Best Local Similarity 48.2%; Pred. No. 0.023;
Matches 134; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

QY 444 tccgtatgtctcaagaattatcatatctctcaagaagctgagacctgcaaacataaagat 503
      ||||| || || |||| | | | | | | | | | | | | | | | | | | | | | | |
DB 1778 tccgtgttcattcgggaatcgtgacaactgtttatgtgtatcttacaataatcagct 1837
      ||||| || || |||| | | | | | | | | | | | | | | | | | | | | | | |
QY 504 tacatccatccatcatatgtcttcagaggaactgaataagccttaactatgtatctcag 563
      ||||| || || |||| | | | | | | | | | | | | | | | | | | | | | | |
DB 1838 ttctgctctatctccgaagaatagttacccgttctcttaccgaactatattggg 1897
      ||||| || || |||| | | | | | | | | | | | | | | | | | | | | | | |
QY 564 tcatacagaataactctcgtgaagccgggtgttttgaagatcttcacagactagaatg 623
      ||||| || || |||| | | | | | | | | | | | | | | | | | | | | | | |
DB 1898 taataactctcttaatgagctctatctcgttcatattgggaatctgaaacaactgtttat 1957
      ||||| || || |||| | | | | | | | | | | | | | | | | | | | | | | |
QY 624 gctgataatgaagaatacctcagtcgaactcccccacaactttatgagactaa 683
      ||||| || || |||| | | | | | | | | | | | | | | | | | | | | | | |
DB 1958 gtgtatcttcaataatcagcttctcgtctatcttccctgaagaataagttacctgag 2017
      ||||| || || |||| | | | | | | | | | | | | | | | | | | | | | | |
QY 684 ttctctatctcttagtctcgtatgaataagctcctca 721
      ||||| || || |||| | | | | | | | | | | | | | | | | | | | | | | |
DB 2018 ttctcttactgaactatttttgggtataaactctcta 2055
      ||||| || || |||| | | | | | | | | | | | | | | | | | | | | | | |

RESULT 14
US-09-180-439-7
; Sequence 7, Application US/09180439
; Patent No. 6225532
; GENERAL INFORMATION:
; APPLICANT: Dixon, Mark S
; APPLICANT: Hatzixanthis, Kostas
; APPLICANT: Jones, David A
; APPLICANT: Jones, Jonathan DG
; TITLE OF INVENTION: Plant pathogen resistance genes and uses thereof
; FILE REFERENCE: 620 - 53
; CURRENT APPLICATION NUMBER: US/09/180,439
; CURRENT FILING DATE: 1998-12-06
; EARLIER APPLICATION NUMBER: PCT/GB97/01249
; EARLIER FILING DATE: 1997-05-08
; EARLIER APPLICATION NUMBER: GB 9609681.3
; EARLIER FILING DATE: 1996-05-09
; EARLIER APPLICATION NUMBER: GB 9619924.5
; EARLIER FILING DATE: 1996-09-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 7
; LENGTH: 4123
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
US-09-180-439-7

Query Match      1.3%; Score 47.6; DB 4; Length 4123;
Best Local Similarity 48.2%; Pred. No. 0.023;
Matches 134; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

QY 444 tccgtatgtctcaagaattatcatatctctcaagaagctgagacctgcaaacataaagat 503
      ||||| || || |||| | | | | | | | | | | | | | | | | | | | | | | |
DB 1922 tccgtgttcattcgggaatcgtgacaactgtttatgtgtatcttacaataatcagct 1981
      ||||| || || |||| | | | | | | | | | | | | | | | | | | | | | | |
QY 504 tacatccatccatcatatgtcttcagaggaactgaataagccttaactatgtatctcag 563
      ||||| || || |||| | | | | | | | | | | | | | | | | | | | | | | |
DB 1982 ttctgctctatctccgaagaatagttacccgttctcttaccgaactatattggg 2041
      ||||| || || |||| | | | | | | | | | | | | | | | | | | | | | | |

```

OY 564 tcaatacgaataacctccgaaagcggtgttttgaagaattctccaagactgaatg 623
| | | | |
DB 2042 taataaacctctaagtcttatcttcgcgtcttaattgggaatccgaacaactgttat 2101

OY 624 gctgataattgaagaataacacactcaagtcgaattccccacaacatttaagactaaa 683
| | | | |
DB 2102 gtgtactttaacataatacgaacttcctgcgtctattctcgaagaataaggttactcgtg 2161

OY 684 ttctctattctcttagtcctgatgaataacgctccta 721
| | | | |
DB 2162 ttctctactgaactatttttggytaataactctcta 2199

RESULT 15
US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEFFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232.463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935.313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZgpt-F1s
US-08-232-463-14

[illegible]

Search completed: September 5, 2002, 08:16:06
Job time: 19001 sec

	Query Match	Similarity	Score	DB	Length
Best Local	1.3%	4.2%	186	131	7218
Matches	14	Conservative	186	Mismatches	131
				Indels	0
				Gaps	0
Oy	2964	ttaatgaacatcatcatatggaatggaatgagatgatgacgacgaagaatg	3023		
Db	1459	TTTAAAGACATGAAATAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRR	1400		
Oy	3024	atcagaaaaactagaatgagataaacatttcattatggaactctcgaataaatcc	3083		

...

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2002, 05:27:20 ; Search time 61.35 Seconds

(without alignments)
1307.177 Million cell updates/sec

Title: US-09-647-067-8

Perfect score: 3834
Sequence: 1 MTSGSVFFYILLFKYFSHG.....FYPCEMSLISOSTPLNLSYS 722Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_032802:*

1: /SID5/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID5/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID5/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID5/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID5/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID5/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID5/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID5/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID5/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID5/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID5/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID5/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID5/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID5/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID5/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID5/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID5/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID5/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID5/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID5/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID5/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3834	100.0	722	20	AA42171	Human IGR7 short f
2	3687.5	96.2	757	20	AA42170	Human IGR7 long fo
3	2091	54.5	396	22	AAE02498	Human CON222 G pro
4	2079	54.2	396	21	AA57286	Human GPCR protein
5	1082.5	28.2	355	22	AAU04370	Human G-protein co
6	948	24.7	188	22	AAAG9945	Human expressed po
7	912	23.8	176	22	AAAG9972	Human expressed po
8	912	23.8	176	22	ABB04062	Human expressed po
9	912	23.8	176	22	ABB10249	Human cdna SEQ ID
10	912	23.8	176	22	AAU18108	Novel human uterin
11	912	23.8	176	22	AAU18344	Human endocrine po

12	912	23.8	176	22	AAU18666	Renal and cardiova
13	912	23.8	176	22	AAU21646	Novel human neopla
14	717.5	18.7	172	22	ABB04105	Human musculoskele
15	650	17.0	140	21	AA841526	Human ORFX ORF1290
16	638	16.6	157	22	AA835407	Human HGRL101 G-pr
17	599.5	15.6	359	22	ABB61216	Drosophila melanog
18	599.5	15.6	359	22	AAU38931	Drosophila G-prote
19	527	13.7	699	14	AAK30517	N-terminal of LH r
20	511.5	13.3	692	14	AAK30503	N-terminal of LH r
21	511.5	13.3	698	14	AAK30505	N-terminal of LH r
22	511	13.3	699	14	AAK30512	N-terminal of LH r
23	492.5	12.8	1300	22	ABB64083	Drosophila melanog
24	490.5	12.8	689	14	AAK30509	N-terminal of LH r
25	489	12.8	699	14	AAK30515	N-terminal of LH r
26	483.5	12.6	695	14	AAK30524	N-terminal of LH r
27	482.5	12.6	695	14	AAK30506	N-terminal of LH r
28	480	12.5	634	14	AAK30520	N-terminal of LH r
29	480	12.5	693	14	AAK30510	N-terminal of LH r
30	479.5	12.5	706	14	AAK30504	N-terminal of LH r
31	479	12.5	685	12	AAK1331	Human luteinizing
32	479	12.5	695	13	AAK27358	FSHR. Homo sapien
33	478.5	12.5	692	11	AAK08038	Rat testicular lut
34	478	12.5	695	14	AAK42082	FSH receptor. Hom
35	478	12.5	695	18	AAW14782	FSH receptor. Hom
36	478	12.5	696	14	AAK30523	N-terminal of LH r
37	477.5	12.5	334	22	ABB60463	Drosophila melanog
38	477.5	12.5	334	22	AAU8929	Drosophila G-prote
39	476.5	12.4	907	20	AAW3889	Human HG38 protein
40	476.5	12.4	907	21	AAV90682	Human G protein-co
41	476.5	12.4	907	21	AAV90687	Human mutant G pro
42	476.5	12.4	951	20	AAV42168	Human IGR4 protein
43	474.5	12.4	700	14	AAK30516	N-terminal of LH r
44	470.5	12.3	695	14	AAK30525	N-terminal of LH r
45	470.5	12.3	700	11	AAK08037	Rat ovarian lutein

ALIGNMENTS

RESULT 1

AA42171

AA42171 standard; Protein; 722 AA.

AC AAV42171;

DE 20-DEC-1999 (first entry)

XX Human IGR7 short form protein sequence.

XX Human IGR4; IGR5; IGR7; G-protein coupled receptor; gene therapy;

KW extracellular leucine rich repeat region; mapping; identification.

XX Homo sapiens.

OS

XX W09948921-AA.

PN

XX 30-SEP-1999.

PD

XX 25-MAR-1999; 99WO-US06573.

PF

XX 26-MAR-1998; 98US-0079501.

PR

XX (STRD) UNIV LELAND STANFORD JUNIOR.

PA (ORGA) ORGANON NV.

PA Hsueh AJW, Hsu SY, Liang S, Van Der Spek PJ;

PI WPI: 1999-591074/50.

XX N-PSDB; AA225346.

DR

XX New G-protein coupled receptors, useful for identifying their own

PT ligands -

XX

PS Claim 2; Fig 4; 54pp; English.

CC The present sequence represents the human G-protein coupled receptor
 CC having extracellular leucine rich repeat regions, designated LGR7 short
 CC form. The LGR4, LGR5 and LGR7 proteins are used to identify ligands for
 CC the receptor. The polypeptides and/or polynucleotides are also useful
 CC for homologous or related genes, producing compositions that modulate
 CC the expression or function of the receptors, gene therapy, mapping
 CC functional regions of the receptors, studying associated physiological
 CC pathways, in vivo prophylactic and therapeutic purposes, as immunogens
 CC for producing antibodies, and for identifying biologically active
 CC agents. The polypeptides contain a G-protein coupled seven
 CC transmembrane region and a leucine rich repeat extracellular domain.
 CC These regions capture and facilitate optimal orientation of its ligand.
 CC The proteins are also expressed in diverse tissues.

XX Sequence 722 AA:

Query Match 100.0%; Score 3834; DB 20; Length 722;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTSQSVFFYLLIEGKFRSHGGGQDVKCSLGYPGCGNTTKCLPOLLHCNGVDDCGNQADE 60
 DB 1 mtsqsvffyllifgkyfshgsgdvkcslygfpcgnttkcpqlhhngyvddcgngaded 60
 QY 61 NCVVVLCQCGSLPGLLELDMMKPFETSPVSNSNTAMSLQNNLIRKLPDPCKNYHDLOKL 120
 DB 61 ncvvvlcqcgsllpgleldmmkpftspvsnsnvtamslqnnlirklpdpcknyhdlokl 120
 QY 121 DLQNNKTSISIVAFRGINSITLTKLYLSHNRITFLKPGVFEDLHRLLEWLIIEDNHLSTRISP 180
 DB 121 dlqnnktsisivafrginsitltklylsnhnrifltpkpgvfedlhrllewlilednhlstrisp 180
 QY 181 PTFYGLNSLILVLMNNVTRLPDCKPLQGMPLRLWLDLBNHINLRLNTFTISCSNWLYV 240
 DB 181 ptfyglnsllilvlmnnvtrlpdckplqgmplrlwldlbnhlnrlntftiscsnwlyv 240
 QY 241 LVMRKKNITNHNENTFAPLQKIDELDLSNKTENLPLRFKDKELISQINLSTVPIQIKQ 300
 DB 241 lvmrkknitnhnentfaplkideldlsnktenlplrfkdkelislqinlstvpiqiqk 300
 QY 301 ANQFDYLVKLSLSLEGIEISNTIQRMFRPLMNLSHYFRKKFOYGYAPHYRSCKPTDG 360
 DB 301 anqfdylvklslslegieisntiqrmfrplmnlshyfrkkfoygyaphyrscckptdg 360
 QY 361 ISSLENLASTIQRFVFWVVSAYTCFENIFVTCMRPTIRSENKLYAMSIISLCCADCLMG 420
 DB 361 isslenlastiqrfvfwvvsaytcfenifvcmrptirsenklyamsiislccadclmg 420
 QY 421 IYLFVIGCFPLKFGGEYFNKAQOLMESTHCOLVSLAIIESTEVSVLLTFTLEKTYCIY 480
 DB 421 iylfvigcfplkfggeyfnkaolmesthcolvslaiiestevsvlltftlektyciy 480
 QY 481 YPRCVRPGKCRITVYLLIMITGFIVAITPLSNKEEFKNYGTNGVCFPLHSEDTESIG 540
 DB 481 yprcvrpgkcrityvllimitgfiivatplsnkeefknnygtngvcfplhsedestig 540
 QY 541 AQIYSVAIFGINNAARIIIVFSGSMFYSVHOSATRTETIRNOKKEMILAKREFITVF 600
 DB 541 aqiyvaifginnaariiivfsgsmfyshosatrtetirnokkemilakrefitvf 600
 QY 601 TDALCWFPIFVWFPLSLQVEIRPCTISWVVFILPINSALNPILLYLTTRPEKEMHRE 660
 DB 601 tdcalfwfpifvwfplslqveirpctiswvfwilpinsalnplillylttrpekemhrrf 660
 QY 661 WYNYRORRKSMDSGQKTYAPSFITWEMVPLQEMPPELMKPDLEFTYPCEMSLSISOSTRLNS 720
 DB 661 wyynyrrorksmdsgqkyapspfitwemvplqemppeilmkpdleftypcemslisostrlns 720
 QY 721 YS 722
 DB 721 YS 722

DB 721 ys 722

RESULT 2

ID AAY42170 standard; Protein; 757 AA.

AC AAY42170;

DT 20-DEC-1999 (first entry)

XX Human LGR7 long form protein sequence.

DE Human: LGR4; LGR5; LGR7; G-protein coupled receptor; gene therapy;

KW extracellular leucine rich repeat region; mapping; identification.

XX Homo sapiens.

PN WO9448921-A1.

PD 30-SEP-1999.

PF 25-MAR-1999; 99MO-US06573.

PR 26-MAR-1998; 98US-0079501.

PA (STRD) UNIV LELAND STANFORD JUNIOR.

XX Hsueh AJW, Hsu SY, Liang S, Van Der Spek PJ;

XX WPI; 1999-591074/50.

DR N-PSDB; AA225345.

PT New G-protein coupled receptors, useful for identifying their own

PS ligands -

XX Claim 2; Fig 3; 54pp; English.

CC The present sequence represents the human G-protein coupled receptor
 CC having extracellular leucine rich repeat regions, designated LGR7 long
 CC form. The LGR4, LGR5 and LGR7 proteins are used to identify ligands for
 CC the receptor. The polypeptides and/or polynucleotides are also useful
 CC for homologous or related genes, producing compositions that modulate
 CC the expression or function of the receptors, gene therapy, mapping
 CC functional regions of the receptors, studying associated physiological
 CC pathways, in vivo prophylactic and therapeutic purposes, as immunogens
 CC for producing antibodies, and for identifying biologically active
 CC agents. The polypeptides contain a G-protein coupled seven
 CC transmembrane region and a leucine rich repeat extracellular domain.
 CC These regions capture and facilitate optimal orientation of its ligand.
 CC The proteins are also expressed in diverse tissues.

XX Sequence 757 AA:

Query Match 96.2%; Score 3687.5; DB 20; Length 757;
 Best Local Similarity 93.0%; Pred. No. 0;
 Matches 707; Conservative 2; Mismatches 10; Indels 41; Gaps 3;

QY 1 MTSQSVFFYLLIEGKFRSHGGGQDVKCSLGYPGCGNTTKCLPOLLHCNGVDDCGNQADE 60
 DB 1 mtsqsvffyllifgkyfshgsgdvkcslygfpcgnttkcpqlhhngyvddcgngaded 60
 QY 61 NC-----VVVLCQCGSLPGLLELDMMKRP- 82
 DB 61 ncgdingwsmqfdkyfasykmtsqypfeecpcclvgsypvqcldq--gldcdetn 117
 QY 83 FTSVPVSNSNTAMSLQNNLIRKLPDCKFNKYHDLOKLDQNNKTSISISYAFRGINSLT 142
 DB 118 ltravpsvsnvtamslqnnlirklpdpckfnkyndloqlqklygmknktsisisyafrginslt 177
 QY 143 KLYLSHNRITFLKPGVFEDLHRLLEWLIIEDNHLSTRISPPTFYGLNSLILVLMNNVTRLP 202

```

Db 178 klyshmrifflkpyfedhrlewlileghnhsrispfcfyglnsllilvlnmvtlrl 237
QY 203 PDKPLOCOMPRCLMDLEGNHINLNLNLPISCSNLTVLVNRKKNINHLNENFAPLQKL 262
Db 238 pdkplocgmprrlmlwldlegnhlnlrltfscsnltvlvnrkknlnhlneftapqlk 297
QY 263 DELDLSGNKTEENPLIFKRLKELSOENLSTYNPORKIOANOFDVLVLRKSLSEGTISN 322
Db 298 deldlsgnktemprllfkdlkeisqlnlsympqlkldanqfolyvklkslsleglsn 357
QY 323 IQQMRERPLNLSHIYKPKQCYGAPRHVSCKPNTDGISSLENLASIIQRFVWVYSA 382
Db 358 lqgrmfprlmnlshlyfkkygcygaphrvsckpntdgiesslenllasllqrvfwvysa 417
QY 383 VTCEGNITVICMPRIYRSENNKLAMSIISCCADCLMGITLVIGFDLFRGEYKHAQ 442
Db 418 vtcfgnlfvicmrylrseknlyamsllsccadclmglylfvlgfdlfrgeynkhaq 477
QY 443 LMWESPHCOLVGLSLISTEVSVLLLFELTEKYICIVYPRCVRPGRKCTITVLLIWI 502
Db 478 lmweshcolvgslslstevsvlllfeltekylcivypfrcvrgpkrctltvliliwi 537
QY 503 TCFIVAFIPLSNKEFEKNYGTNGVCPFLHSEDTESIGAOIYSVAIFLGINLAFTITVE 562
Db 538 tcfivafiplsnkefknnygtngvcfplhsestesigadlysvaiflglnlaaftitve 597
QY 563 SYGSMFYSVQSATITATEIRNOVKKEMLAKRPFYITFDALCMIPIFYVKSLSLOVEI 622
Db 598 sygsmfysvqsaltateirnvkkelmlakrffifvftdalcmipifvyklsllqvel 657
QY 623 PGRTTSMVVFILPINSALNPILYTLFTRPPEKEMIHREWYNYRORKMDSGOKTYPSF 682
Db 658 pgrttsmvvfilpinaalnpllytlctrpfkemihrtwynytrksmdskgqktypsf 717
QY 683 IWEMMPLOEMPELMKPDLETFYPCENSLISQSTRLSYS 722
Db 718 iwemmplogempelmkpdltfypcensllsqstrlnsys 757

RESULT 3
AAE02498
ID AAE02498 standard; Protein; 396 AA.
XX
AC AAE02498;
XX
DE 10-AUG-2001 (first entry)
XX
DE Human CON222 G protein-coupled receptor protein.
XX
KW Human; G protein-coupled receptor; GPCR; CON222 protein; schizophrenia;
KW neuroleptic; nootropic; neuroprotective; bipolar disease; psychotropic;
KW neurological disorder; psychiatric disease; neurosis; anxiety; neuritis;
KW attention deficit hyperactivity disorder; neuroshtia; senile dementia;
KW affective disorder; neuropathy; Alzheimer's disease; Parkinson's disease;
KW depression; migraine; genetic screening.
XX
OS Homo sapiens.
XX
XX
FH Key
FT location/Qualifiers
FT 42..65
FT /label= Transmembrane_domain_(1TW)
FT 66..78
FT /label= Intracellular_domain
FT /note= "First IC loop"
FT 79..103
FT Domain
FT /label= Transmembrane_domain_(2TW)
FT 104..124
FT /label= Extracellular_domain
FT /note= "First EC loop"
FT 125..156
FT Domain
FT /label= Transmembrane_domain_(3TW)
FT 157..166
FT Domain

```

```

FT /label= Intracellular_domain
FT /note= "Second IC loop"
FT 167..188
FT Domain
FT /label= Transmembrane_domain_(4TW)
FT 189..216
FT Domain
FT /label= Extracellular_domain
FT /note= "Second EC loop"
FT 217..241
FT Domain
FT /label= Transmembrane_domain_(5TW)
FT 242..267
FT /label= Intracellular_domain
FT /note= "Third IC loop"
FT 268..290
FT Domain
FT /label= Transmembrane_domain_(6TW)
FT 291..300
FT /label= Extracellular_domain
FT /note= "Third EC loop"
FT 301..320
FT Domain
FT /label= Transmembrane_domain_(7TW)

W0200131014-A2.
XX
XX 03-MAY-2001.
XX
XX 27-OCT-2000; 2000WO-US29601.
XX
XX 27-OCT-1999; 99US-0427653.
XX 27-OCT-1999; 99US-0427859.
XX 27-OCT-1999; 99US-0428020.
XX 27-OCT-1999; 99US-0428114.
XX 28-OCT-1999; 99US-0429517.
XX 28-OCT-1999; 99US-0429555.
XX 28-OCT-1999; 99US-0429676.
XX 28-OCT-1999; 99US-0429695.
XX 28-OCT-1999; 99US-0454399.
XX 12-JAN-2000; 2000US-0481794.
XX
XX (PHAA ) PHARMACIA & UPJOHN CO.
XX
XX Vogel G, Wood LS, Merchant K;
XX
XX WPI; 2001-328653/34.
XX
XX N-PSDB; AAD06507.
XX
XX
XX Seven transmembrane receptor polypeptides and polynucleotides, useful
XX for treating neurological or psychiatric disorders, e.g. schizophrenia,
XX as well as for identifying compounds useful for treating schizophrenia
XX
XX
XX Claim 1; Page 15-16; 215pp; English.
XX
XX The invention relates to human G protein-coupled receptor (GPCR) and
XX their corresponding DNA molecules. GPCR is also referred as seven
XX transmembrane receptor. G protein-coupled receptor protein is useful for
XX treating neurological disorder, particularly schizophrenia. GPCR protein
XX is also useful for identifying compounds useful for treating other
XX schizophrenia. These compounds are also useful for treating other
XX neurological and psychiatric diseases, e.g. depression, anxiety, bipolar
XX disease, affective disorders, attention deficit hyperactivity disorder/
XX attention deficit disorder, epilepsy, neuritis, neuroshtia, neuropathy,
XX neurosis, Alzheimer's disease, Parkinson's disease, migraine and senile
XX dementia. The invention also provides genetic screening procedures that
XX entail analysing a person's genome with respect to GPCR. The vectors are
XX useful for the recombinant production of the GPCR's. The present sequence
XX is human CON222 G protein-coupled receptor (GPCR) protein.
XX
XX
XX Sequence 396 AA;

```

Query Match 54.5%; Score 2091; DB 22; Length 396;
 Best Local Similarity 99.7%; Pred. No. 7e-195;
 Matches 395; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY	327	MEPRMLNLSHIEYKFKFOYCGYAPHVNSCKRNTDGSISLEMLNLSIIORVWVWVSAVTC	386
Db	1	mfrplvnlsihyfkfkfcgcyaphvnsckpntdgsislenllasilqrfvfwvwsavtc	60
QY	387	GNIFVICMRPYIRENKLTYMSITSLCCACDCLNGITYFVVGPGPLKRGSEYKNAHQAQME	446
Db	61	gnltfvlcmrpylrseemlyamsllslccacclnglyltfvtvggfdlkrtyeynhaqame	120
QY	447	STHCOLVGSLLSLSTEVSVLLNFTPLTELEKYICIVYPRRCRVRPCKRTITVLILMTWTGFI	506
Db	121	sthcolvgslslstevsvllftltlekyicivprcrrpckrtltvllilwtgfl	180
QY	507	VAFIPLSKKEEFKNYVGTNGVCPFLHSEDBESICAOQIYVAIFLGIMLAFIITVSYGS	566
Db	181	vafiplskkefkfnygvngvcfplhsedbeslgaqlyvaiflgilaaftiaivfsygs	240
QY	567	MEFVSHOSATATELRNOKKEMILARFPFIYTTDLQWPIRVVFELSULOVEIGTI	626
Db	241	mefvshosatlattelrnokkemilarkfflftvltldacwplrvvfelslvgelpgtl	3000
QY	627	TSWVVFILPINSALNPILYLTITRPRKEMIRHWYVNRORSMDSKGOKTYAPSFITWE	686
Db	301	tswvvtlflpnsalnpllylttrprfkemhrtwryvnrqskmdskgoktyapsfittwe	3606
QY	687	MMPLQEMPPELMKPDLETTYPCEMSLISOSTRFLNYS	722
Db	361	mmplqemppeilmkpdlfttypcemslisosttrflnys	396

CC	AA57286	standard. Protein; 396 AA.
XX	AA57286;	
XX	05-JUN-2000	(first entry)
DE	Human GPCR protein (HGPRP) sequence (clone ID 2488822).	
XX	Human; G protein coupled protein receptor; HGPRP; cell proliferation;	
KW	neurological; immune disorder; cytostatic; anti-arteriosclerotic;	
KW	anti-atherosclerotic; hepatotropic; antiinflammatory; virucide; leukemia;	
KW	immunomodulatory; anemia; asthma; gastrointestinal; anti-epileptic;	
KW	anti-Alzheimer's; anti-Parkinsonian; gene therapy.	
OS	Homo sapiens.	
XX	WO200015793-A2.	
XX	23-MAR-2000.	
XX	17-SEP-1999;	99WO-US20958.
XX	17-SEP-1998;	98US-0156513.
XX	(INCY-) INCYTE PHARM INC.	
XX	Bandman O, Lal P, Tang YT, Corley NC, Guegler KU, Gorgone GA;	
PI	Baughn MR;	
XX	WPI: 2000-271432/23.	
DR	N-PSDB; AA290524.	
XX	Human G protein coupled protein receptor peptides useful for the	
PT	prevention, diagnosis and treatment of cell proliferative, neurological	
PT	and immune disorders -	
XX	Claim 1; Page 62-63; 71pp: English.	
XX	The invention provides human G protein coupled protein receptor (HGPRP)	
CC	polypeptides and polynucleotides encoding them. The polypeptides can be	
CC	produced by standard recombinant methodology. The polynucleotides and	
CC	polypeptides may be used in the prevention, treatment and diagnosis of	

CC diseases associated with their inappropriate expression. Diseases that
CC can be treated are cell proliferative disorders (e.g. arteriosclerosis
CC atherosclerosis and hepatitis), cancers (e.g. leukemia, melanomas and
CC adenocarcinoma), immune disorders (e.g. anemia, asthma and Crohn's
CC disease) and neurological disorders (e.g. epilepsy, Alzheimer's disease
CC and Parkinson's disease). The anti-hGPCR antibodies may also be used as
CC diagnostic agents for detecting the presence of GPCR polypeptides in
CC samples (e.g. by enzyme linked immunosorbent assay (ELISA)). Sequences
CC AA57283-288 represent the hGPCR polypeptides.

Query Match	54.28;	Score 2079;	DB 21;	Length 396;
Best Local Similarity	99.2%;	Pred. No. 1e-193;		
Matches 393; Conservative	1;	Mismatches 2;	Indels 0;	Gaps 0;

```

Oy 327 MRRPMLNLSHITYFKKFOYGYAPHVRSCKPNTDGISSLEMLIASI IORVWVWVSANTCF 366
Db 1 mfrplvnishiyfkfgygyaphvrscckpntdgiisslenllasllqrvftvwvwsavtcf 60
Oy 387 GNIFPICMRPYIRSRNKIKYAMSIISLCCADLMGTYFLVGGPDLKRGEXNKHQALME 446
Db 61 gnifpicmrpyirsrncklyamslslccadclmgtyflvvggdlkrtgrcnkhaqlme 120
Oy 447 SNHCOLVSLALISTEYVVLTLTFELTEKYICIIYPERFCVAPGCKRTIYVLLIWMNGFI 506
Db 121 schcolvyslalistevsylltftlekycivypfrcvrgpcrctlvllllwlgfl 180
Oy 507 VAFIPLSNKEFPKNYGTNGVCFPLHSDTESIGAOIYSAVIFGJNLMAFIITVESYGS 566
Db 181 vafiplsnkeffknuyapnvcvfpshsdteslgaqysvalfignllaaftllvfyysgs 240
Oy 567 MFTSYHQSAITRTLRNOYKKEMLAKRFFPIVLPDLCWPIRVVFLSLLOVEIGTI 626
Db 241 mftsyhsqalitecltrnvqvkemllakffivtfidalcwpiplrvvflsllyveipgtl 3000
Oy 627 TSMVYIFLPIINSALNPILYTLTRPFEMIHREYVNIORRKSNDKSGOQYTABSFIVE 686
Db 301 tsmvyiflpiinsalnpllytltrpfemihrfeynyqrkmsdskqcyapbsflwe 3600
Oy 687 KMPLOEMPELMKPDLFYPCMSLISOSTRLNYS 722
Db 361 kmplgempeelmkpdlfycpmslissstrlnys 396

```

PR	23-DEC-1999;	99US-0171902.
PR	23-DEC-1999;	99US-0171901.
PR	23-DEC-1999;	99US-0171900.
PR	17-NOV-1999;	99US-0166369.
PR	17-NOV-1999;	99US-0166088.
XX		
XX	16-NOV-2000;	2000MO-US31509.
XX	PE	
XX	25-MAY-2001.	
XX	PD	
XX	PN	WO200136471-A2.
XX	OS	Homo sapiens.
XX	KM	Inverse agonist; lung cancer.
XX	KM	Human; G-protein coupled receptor; GPCR; hRUP16; agonist.
XX	DE	Human G-protein coupled receptor, hRUP16.
XX	DE	
XX	DT	23-OCT-2001 (first entry)
XX	DT	
XX	AC	AAU04370;
XX	ID	AAU04370 standard; Protein; 355 AA.
RESULT	5	

PR 11-FEB-2000; 2000US-0181749.
PR 14-MAR-2000; 2000US-0189258.
PR 14-MAR-2000; 2000US-0189259.
PR 10-APR-2000; 2000US-0195898.
PR 10-APR-2000; 2000US-0195899.
PR 28-APR-2000; 2000US-0196078.
PR 12-MAY-2000; 2000US-0203630.
PR 12-JUN-2000; 2000US-0210741.
PR 12-JUN-2000; 2000US-0210982.
PR 21-AUG-2000; 2000US-0226760.
PR 26-SEP-2000; 2000US-0235418.
PR 26-SEP-2000; 2000US-0235779.
PR 20-OCT-2000; 2000US-0242332.
PR 20-OCT-2000; 2000US-0242343.
XX
XX (AREN-) ARENA PHARM INC.
XX
XX Chen R, Dang HT, Lowitz KP;
XX
XX WPI: 2001-355616/37.
XX
XX N-PSDB; AAS07943.
XX
XX Endogenous and non-endogenous versions of human G-protein coupled
PT receptors for direct identification of candidate compounds as agonists,
PT inverse agonists or partial agonists for use as therapeutic agents -
PS Claim 33; Page 105-106; 160pp; English.
XX
XX The sequence represents a human G-protein coupled receptor (GPCR),
CC hRup16. The endogenous and non-endogenous, constitutively activated
CC versions of human G-protein coupled receptors (GPCR), are useful for
CC direct identification of candidate compounds as receptor agonists,
CC inverse agonists or partial agonists having applicability as therapeutic
CC agents for treating diseases related to GPCR, e.g. lung cancer.
CC Non-endogenous version of human GPCRs are also utilized in research
CC settings and in vitro and in vivo system, incorporating GPCRs can be
CC utilized to elucidate and understand the roles these receptors
CC play in the human condition, both normal and diseased.
XX
XX Sequence 355 AA;
SQ

Query Match 28.2%; Score 1082.5; DB 22; Length 355;
Best Local Similarity 59.8%; Pred. No. 1.le-96;
Matches 198; Conservative 67; Mismatches 65; Indels 1; Gaps 1;

QY 356 PNTDGISSLENTLSTIQRFEVWVSAVTCFNIPIVCMKRPYRSFNKLYAMSIISLCCA 415
DB 2 pldglsstfedlannllirfivvaiafclfnllygmrsfikaentthamsikilcca 61
QY 416 DCLMGIVLYFVGFDLKEFGEVYKHAQIMESTHCQLVGSIALSTEVSVLLTFELTERK 475
DB 62 dclmgvlyffvgfdlkyrgyqkyaallwmesvqcrflmglamlslevsvllltytlek 121
QY 476 YICIVYPPFCVCRKQRTITVILIMWINGFIVAFPLSKKEFKNYGNGVCFPLHSD 535
DB 122 flivvfpfnslpfgkrtqvslilclmagflavlpfwnkdyfngfygknvcfpilyydg 181
QY 536 TESIGAQIVSAFLGINLAFILIVFSGMFSYSHQSAITETIRNOVKKEMILAKKF 595
DB 182 tedisgkysjlfifgnvlnlafllivfsytlmfscqkalqetternofgrevaanf 241
QY 596 FPIVETDALCWPPIVYKFLSLQVEIPGITSWVIFILPINSALINPLIYLLTRPEKE 655
DB 242 fpiivetdalcwppivvkvkflslqlveipgtmswvifilpfnalsainpllyllttnfk 301
QY 656 MHHREFYNYRQRKSMDSKQKTYAPSEIWE 686
DB 302 kikqllhkn-qtrkslfrkklkksisctsiwle 331

AAM99945
ID AAM99945 standard; Protein; 188 AA.
XX
XX AAM99945;
XX
XX 04-JAN-2002 (first entry)
XX
XX Human expressed polypeptide SEQ ID NO 69.
DE
XX Human; nootropic; neuroprotective; cyostatic; dermatological; virocidic;
KW immunosuppressive; antinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antiskinning; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antinflammatory;
KW antiallergic; antidiabetic; antilicer; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine.
XX
XX Homo sapiens.
XX
XX WO20015387-A1.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01310.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 04-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0225759.
PR 22-AUG-2000; 2000US-0226881.
PR 22-AUG-2000; 2000US-0226886.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.

RESULT 7
ID AAM9972 standard; Protein: 176 AA.
XX AAM9972;
AC
XX
DT 04-JAN-2002 (first entry)
XX
DE Human expressed polypeptide SEQ ID NO 96.
XX
KW Human; nootropic; neuroprotective; cytosolic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;
KW antiparkinsonian; antisticking; antianemic; antiarthritic; cancer;
KW antineumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antilicer; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine.
XX
OS Homo sapiens.
XX
FN W020015387-A1.
PD
XX 02-AUG-2001.
PE 17-JAN-2001; 2001WO-US01310.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 23-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.

PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.

PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249267.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.

PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251968.
PR 05-DEC-2000; 2000US-0251969.
PR 06-DEC-2000; 2000US-0251479.
PR 06-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251858.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-451937/48.
DR N-PSDB; AAL35644.
DR
XX
PT Isolated polypeptide for treating, preventing and/or prognosing
PT disorders related to the musculoskeletal system including
PT musculoskeletal cancers and also for testing and detection e.g.
PT diagnosis -
XX
PS Claim 11; SEQ ID NO 2009; 781pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (AAL34669-AAL37666) and proteins
CC (ABB03087-ABB04109) associated with the musculoskeletal system useful
CC for preventing, treating or ameliorating medical conditions e.g. by
CC protein or gene therapy. The genes are isolated from a range of human
CC tissues disclosed in the specification. The nucleic acids, proteins,
CC antibodies and (ant)agonists are useful in the diagnosis, treatment,
CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and
CC other cancers of the adrenal gland, bone, bone marrow, breast,
CC gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC (c) cardiovascular disorders such as myocardial ischemias; (d) wound
CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
CC and (f) infectious diseases such as viral, bacterial, fungal and
CC parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
CC
XX
SQ Sequence 176 AA;

Query Match 23.88; Score 912; DB 22; Length 176;
Best Local Similarity 99.48; Pred. No. 1.7e-80;
Matches 175; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 541 AQTVAIFLCINLAFTIIFVSGSMFYSHQSAITATEIRNQYKEMITAKRPFIVF 600
DB 1 aqlysvaiflginlaaifilvfygsmfysbgsaitetelnqykemilakiffivf 60
QY 601 TDALCWPIFYVKELSLQVEIPIGTTSWVFIILPINSALNPILYTLTTPRFKEMIRF 660
DB 61 tdaicwpiifvkvklsllqveipgltswwvifilpinsalnpllytltpfkemlrif 120
QY 661 WYNTRQKSNDSKQOKTAPSFIVEMKPIIDMPPELKKPDLFTYPCMSLISQST 716
DB 121 wyntrqksmdskqktyapsfiwemwplqempeimkpdltfycemsllsqst 176

RESULT 9
ABBI0249
ID ABBI0249 standard; Protein; 176 AA.
XX
AC ABBI0249;
XX

DT 10-JAN-2002 (first entry)
XX Human CDNA SEQ ID NO: 557.
XX
KW Human; gene therapy; neural disorder; immune system disorder;
muscular disorder; reproductive disorder; gastrointestinal disorder;
pulmonary disorder; cardiovascular disorder; renal disorder;
proliferative disorder; inflammation.
XX
OS Homo sapiens.
XX
FN WO200154474-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01349.
XX
PR 31-JAN-2000; 2000US-179065P.
PR 04-FEB-2000; 2000US-180628P.
PR 24-FEB-2000; 2000US-184664P.
PR 02-MAR-2000; 2000US-186350P.
PR 16-MAR-2000; 2000US-189874P.
PR 17-MAR-2000; 2000US-190076P.
PR 18-APR-2000; 2000US-198123P.
PR 19-MAY-2000; 2000US-205515P.
PR 07-JUN-2000; 2000US-209467P.
PR 28-JUN-2000; 2000US-214886P.
PR 30-JUN-2000; 2000US-215135P.
PR 07-JUL-2000; 2000US-216647P.
PR 07-JUL-2000; 2000US-216880P.
PR 11-JUL-2000; 2000US-217487P.
PR 11-JUL-2000; 2000US-217496P.
PR 14-JUL-2000; 2000US-218290P.
PR 26-JUL-2000; 2000US-220963P.
PR 26-JUL-2000; 2000US-220964P.
PR 14-AUG-2000; 2000US-224518P.
PR 14-AUG-2000; 2000US-224519P.
PR 14-AUG-2000; 2000US-225213P.
PR 14-AUG-2000; 2000US-225214P.
PR 14-AUG-2000; 2000US-225267P.
PR 14-AUG-2000; 2000US-225267P.
PR 14-AUG-2000; 2000US-225268P.
PR 14-AUG-2000; 2000US-225270P.
PR 14-AUG-2000; 2000US-225447P.
PR 14-AUG-2000; 2000US-225757P.
PR 14-AUG-2000; 2000US-225757P.
PR 14-AUG-2000; 2000US-225759P.
PR 14-AUG-2000; 2000US-226279P.
PR 18-AUG-2000; 2000US-226681P.
PR 22-AUG-2000; 2000US-226681P.
PR 22-AUG-2000; 2000US-226868P.
PR 22-AUG-2000; 2000US-227182P.
PR 23-AUG-2000; 2000US-227009P.
PR 30-AUG-2000; 2000US-228924P.
PR 01-SEP-2000; 2000US-229287P.
PR 01-SEP-2000; 2000US-229343P.
PR 01-SEP-2000; 2000US-229344P.
PR 01-SEP-2000; 2000US-229345P.
PR 05-SEP-2000; 2000US-229509P.
PR 05-SEP-2000; 2000US-229513P.
PR 06-SEP-2000; 2000US-229547P.
PR 06-SEP-2000; 2000US-230438P.
PR 08-SEP-2000; 2000US-231242P.
PR 08-SEP-2000; 2000US-231243P.
PR 08-SEP-2000; 2000US-231244P.
PR 08-SEP-2000; 2000US-231244P.
PR 08-SEP-2000; 2000US-231413P.
PR 08-SEP-2000; 2000US-231414P.
PR 08-SEP-2000; 2000US-232088P.
PR 08-SEP-2000; 2000US-232081P.
PR 12-SEP-2000; 2000US-231968P.
PR 14-SEP-2000; 2000US-232397P.
PR 14-SEP-2000; 2000US-232398P.
PR 14-SEP-2000; 2000US-232399P.
PR 14-SEP-2000; 2000US-232400P.

PR 14-SEP-2000; 2000US-232401P.
PR 14-SEP-2000; 2000US-233063P.
PR 14-SEP-2000; 2000US-233064P.
PR 14-SEP-2000; 2000US-233065P.
PR 21-SEP-2000; 2000US-234223P.
PR 21-SEP-2000; 2000US-234274P.
PR 25-SEP-2000; 2000US-234997P.
PR 25-SEP-2000; 2000US-234998P.
PR 26-SEP-2000; 2000US-235844P.
PR 27-SEP-2000; 2000US-235834P.
PR 27-SEP-2000; 2000US-235836P.
PR 29-SEP-2000; 2000US-236327P.
PR 29-SEP-2000; 2000US-236367P.
PR 29-SEP-2000; 2000US-236368P.
PR 29-SEP-2000; 2000US-236369P.
PR 29-SEP-2000; 2000US-236570P.
PR 29-SEP-2000; 2000US-236802P.
PR 02-OCT-2000; 2000US-237037P.
PR 02-OCT-2000; 2000US-237038P.
PR 02-OCT-2000; 2000US-237039P.
PR 02-OCT-2000; 2000US-237040P.
PR 13-OCT-2000; 2000US-239335P.
PR 13-OCT-2000; 2000US-239337P.
PR 20-OCT-2000; 2000US-240960P.
PR 20-OCT-2000; 2000US-241221P.
PR 20-OCT-2000; 2000US-241785P.
PR 20-OCT-2000; 2000US-241786P.
PR 20-OCT-2000; 2000US-241787P.
PR 20-OCT-2000; 2000US-241808P.
PR 20-OCT-2000; 2000US-241809P.
PR 20-OCT-2000; 2000US-241826P.
PR 01-NOV-2000; 2000US-244617P.
PR 08-NOV-2000; 2000US-246474P.
PR 08-NOV-2000; 2000US-246475P.
PR 08-NOV-2000; 2000US-246476P.
PR 08-NOV-2000; 2000US-246477P.
PR 08-NOV-2000; 2000US-246478P.
PR 08-NOV-2000; 2000US-246523P.
PR 08-NOV-2000; 2000US-246524P.
PR 08-NOV-2000; 2000US-246525P.
PR 08-NOV-2000; 2000US-246526P.
PR 08-NOV-2000; 2000US-246527P.
PR 08-NOV-2000; 2000US-246528P.
PR 08-NOV-2000; 2000US-246532P.
PR 08-NOV-2000; 2000US-246609P.
PR 08-NOV-2000; 2000US-246610P.
PR 08-NOV-2000; 2000US-246611P.
PR 08-NOV-2000; 2000US-246613P.
PR 17-NOV-2000; 2000US-249207P.
PR 17-NOV-2000; 2000US-249208P.
PR 17-NOV-2000; 2000US-249209P.
PR 17-NOV-2000; 2000US-249210P.
PR 17-NOV-2000; 2000US-249211P.
PR 17-NOV-2000; 2000US-249212P.
PR 17-NOV-2000; 2000US-249213P.
PR 17-NOV-2000; 2000US-249214P.
PR 17-NOV-2000; 2000US-249215P.
PR 17-NOV-2000; 2000US-249216P.
PR 17-NOV-2000; 2000US-249217P.
PR 17-NOV-2000; 2000US-249218P.
PR 17-NOV-2000; 2000US-249244P.
PR 17-NOV-2000; 2000US-249245P.
PR 17-NOV-2000; 2000US-249246P.
PR 17-NOV-2000; 2000US-249256P.
PR 17-NOV-2000; 2000US-249257P.
PR 17-NOV-2000; 2000US-249297P.
PR 17-NOV-2000; 2000US-249300P.
PR 01-DEC-2000; 2000US-250160P.
PR 01-DEC-2000; 2000US-250391P.
PR 05-DEC-2000; 2000US-251030P.
PR 05-DEC-2000; 2000US-251988P.
PR 05-DEC-2000; 2000US-256719P.
PR 06-DEC-2000; 2000US-251479P.

PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246539.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249254.
PR 17-NOV-2000; 2000US-0249255.
PR 17-NOV-2000; 2000US-0249257.
PR 17-NOV-2000; 2000US-0249259.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251968.
PR 05-DEC-2000; 2000US-0251968.
PR 05-DEC-2000; 2000US-0251969.
PR 05-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI: 2001-488777/53.
DR N-PSDB; AAS28950.
XX
XX Isolated polypeptide and nucleic acid molecules for treating,
PT preventing and/or prognosing disorders related to uterine motility
PT e.g. disorders associated with pregnancy and the menstrual cycle -
XX
XX PS Claim 11; SEQ ID No 84; 524pp; English.
XX
XX The present invention relates to the isolation of novel human

CC uterine motility-association polypeptides, and cDNA (AAS28936-AAS28994)
-CC and genomic sequences encoding for these polypeptides.
CC The sequences of the invention are useful in the diagnosis,
CC treatment, prevention and/or prognosis of diseases associated
CC with uterine motility such as pregnancy and labour, and menstrual
CC disorders. The polynucleotide sequences of the invention are also
CC useful in gene therapy. AAU18094-AAU18152 represent novel human
CC uterine motility-association polypeptides.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 176 AA;

Query Match 23.8%; Score 912; DB 22; Length 176;
Best Local Similarity 99.4%; Pred. No. 1.7e-80;
Matches 175; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 541 AQLYSVAIFLGINLAFLIIVSYSGMEYSVHQSATATEIRNOVKREMLAKRPFIVF 600
|||||
Db 1 aqysvaiflginlaaflilivfsgmsmlyshgaitatelnqykemlakrffivf 60
|||||
QY 601 TDALCWIPDFYVKFSLQVEIPEPTITSWVIFILPINSALNPILYLTTPPEKMIHRE 660
|||||
Db 61 tdaicwpldfvkvkflslqlveipgltswwvifilpnsalnpllylttrpfkcmahrf 120
|||||
QY 661 WYNRORKSMDSKGOKTYAPSEFIWEMKPLQEMPELKKPDLFTYPCMSLISQST 716
|||||
Db 121 wynrqksmdskgktyapseflwemwplqempeimkpoltypcemslisqst 176
|||||

RESULT 11
AAU18344
ID AAU18344 standard; Protein: 176 AA.
XX
AC AAU18344;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human endocrine polypeptide SEQ ID No 299.
XX
XX Endocrine protein; human; mouse; rabbit; goat; horse; food additive;
KW cat; dog; chicken; sheep; immunosuppressive; antihypertic; vasotropic;
KW antineumatic; antiproliferative; cytostatic; cardiant; neuroprotective;
KW cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;
KW ophthalmological; vulnerrary; gene therapy; autoimmune disease; neoplasm;
KW hyperproliferative disorder; breast; liver; cardiovascular disorder;
KW cerebrovascular disorder; nervous system disorder; bacterial infection;
KW fungal infection; viral infection; ocular disorder; endocrine disorder;
KW gastrointestinal disorder; renal disorder; respiratory disorder;
KW wound healing; skin aging; organ transplantation; food preservative;
KW tissue regeneration; anti-fertility.
XX
OS Homo sapiens.
XX
OS
XX
PN WO20015364-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01308.
XX
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
XX 24-FEB-2000; 2000US-0184664.
XX 02-MAR-2000; 2000US-0186350.
XX 16-MAR-2000; 2000US-0189874.
XX 17-MAR-2000; 2000US-0190076.
XX 18-APR-2000; 2000US-0198123.
XX 19-MAY-2000; 2000US-0205515.
XX 07-JUN-2000; 2000US-0209467.
XX 28-JUN-2000; 2000US-0214886.
XX 30-JUN-2000; 2000US-0215135.
PR

PR	07-JUL-2000	2000US-0216647
PR	07-JUL-2000	2000US-0216680
PR	11-JUL-2000	2000US-0211487
PR	11-JUL-2000	2000US-0211796
PR	14-JUL-2000	2000US-0218290
PR	26-JUL-2000	2000US-0220963
PR	26-JUL-2000	2000US-0220964
PR	14-AUG-2000	2000US-0224518
PR	14-AUG-2000	2000US-0224519
PR	14-AUG-2000	2000US-0225514
PR	14-AUG-2000	2000US-0225516
PR	14-AUG-2000	2000US-0225758
PR	14-AUG-2000	2000US-0225759
PR	18-AUG-2000	2000US-0226779
PR	22-AUG-2000	2000US-0226681
PR	22-AUG-2000	2000US-0226688
PR	23-AUG-2000	2000US-0227182
PR	23-AUG-2000	2000US-0227209
PR	30-AUG-2000	2000US-0228824
PR	01-SEP-2000	2000US-0229287
PR	01-SEP-2000	2000US-0229343
PR	01-SEP-2000	2000US-0229344
PR	01-SEP-2000	2000US-0229345
PR	05-SEP-2000	2000US-0229509
PR	05-SEP-2000	2000US-0229513
PR	06-SEP-2000	2000US-0230437
PR	06-SEP-2000	2000US-0230438
PR	08-SEP-2000	2000US-0231342
PR	08-SEP-2000	2000US-0231343
PR	08-SEP-2000	2000US-0231344
PR	08-SEP-2000	2000US-0231414
PR	08-SEP-2000	2000US-0231414
PR	08-SEP-2000	2000US-0232080
PR	08-SEP-2000	2000US-0232081
PR	12-SEP-2000	2000US-0231368
PR	14-SEP-2000	2000US-0232397
PR	14-SEP-2000	2000US-0232398
PR	14-SEP-2000	2000US-0232399
PR	14-SEP-2000	2000US-0232401
PR	14-SEP-2000	2000US-0233063
PR	14-SEP-2000	2000US-0233064
PR	14-SEP-2000	2000US-0233065
PR	21-SEP-2000	2000US-0234423
PR	21-SEP-2000	2000US-0234424
PR	25-SEP-2000	2000US-0234897
PR	25-SEP-2000	2000US-0234998
PR	26-SEP-2000	2000US-0234584
PR	27-SEP-2000	2000US-0235834
PR	27-SEP-2000	2000US-0235836
PR	29-SEP-2000	2000US-0236327
PR	29-SEP-2000	2000US-0236367
PR	29-SEP-2000	2000US-0236368
PR	29-SEP-2000	2000US-0236369
PR	02-OCT-2000	2000US-0236370
PR	02-OCT-2000	2000US-0236802
PR	02-OCT-2000	2000US-0237037
PR	02-OCT-2000	2000US-0237038
PR	02-OCT-2000	2000US-0237309
PR	02-OCT-2000	2000US-0237340
PR	13-OCT-2000	2000US-0239935
PR	13-OCT-2000	2000US-0239937
PR	20-OCT-2000	2000US-0240960
PR	20-OCT-2000	2000US-0241121
PR	20-OCT-2000	2000US-0241185
PR	20-OCT-2000	2000US-0241786
PR	20-OCT-2000	2000US-0241787

PR	20-OCT-2000;	2000US-0241808.
PR	20-OCT-2000;	2000US-0241809.
PR	20-OCT-2000;	2000US-0241826.
PR	01-NOV-2000;	2000US-0244617.
PR	08-NOV-2000;	2000US-0246474.
PR	08-NOV-2000;	2000US-0246475.
PR	08-NOV-2000;	2000US-0246476.
PR	08-NOV-2000;	2000US-0246477.
PR	08-NOV-2000;	2000US-0246478.
PR	08-NOV-2000;	2000US-0246523.
PR	08-NOV-2000;	2000US-0246524.
PR	08-NOV-2000;	2000US-0246525.
PR	08-NOV-2000;	2000US-0246526.
PR	08-NOV-2000;	2000US-0246527.
PR	08-NOV-2000;	2000US-0246528.
PR	08-NOV-2000;	2000US-0246532.
PR	08-NOV-2000;	2000US-0246609.
PR	08-NOV-2000;	2000US-0246610.
PR	08-NOV-2000;	2000US-0246611.
PR	08-NOV-2000;	2000US-0246613.
PR	17-NOV-2000;	2000US-0249207.
PR	17-NOV-2000;	2000US-0249208.
PR	17-NOV-2000;	2000US-0249209.
PR	17-NOV-2000;	2000US-0249210.
PR	17-NOV-2000;	2000US-0249211.
PR	17-NOV-2000;	2000US-0249212.
PR	17-NOV-2000;	2000US-0249213.
PR	17-NOV-2000;	2000US-0249214.
PR	17-NOV-2000;	2000US-0249215.
PR	17-NOV-2000;	2000US-0249216.
PR	17-NOV-2000;	2000US-0249217.
PR	17-NOV-2000;	2000US-0249218.
PR	17-NOV-2000;	2000US-0249244.
PR	17-NOV-2000;	2000US-0249245.
PR	17-NOV-2000;	2000US-0249264.
PR	17-NOV-2000;	2000US-0249265.
PR	17-NOV-2000;	2000US-0249267.
PR	17-NOV-2000;	2000US-0249299.
PR	17-NOV-2000;	2000US-0249300.
PR	01-DEC-2000;	2000US-0250160.
PR	01-DEC-2000;	2000US-0250391.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	05-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251989.
PR	08-DEC-2000;	2000US-0251990.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0259678.
XX	(HUMA-) HUMAN GENOME SCI INC.	
PI	Rosen CA, Barash SC, Ruben SM;	
XX	WPI: 2001-451936/48.	
DR	N-PSDB: AAS29573.	
XX		
PT	Isolated polypeptide for treating, preventing and/ or prognosing	
PT	disorders of the endocrine system such as reproductive disorders,	
PT	endocrine cancers and also for testing and detection e.g. diagnosis -	
XX		
PS	Claim 11; SEQ ID NO 299; 604bp; English.	
XX		
CC	Sequences AAU1282-AAU18507 represent endocrine polypeptides of the	
CC	invention. Endocrine polypeptides and their associated polynucleotides	
CC	are useful in the diagnosis, treatment and prevention of various types of	
CC	disorders in e.g. humans, mice, rabbits, goats, cats, dogs,	
CC	chickens or sheep. A pathological condition can be determined by	
CC	determining the presence or absence of a mutation in an endocrine	
CC	polynucleotide. The treatable disorders include autoimmune diseases such	

CC as rheumatoid arthritis, hyperproliferative disorders such as neoplasms
CC of the breast or liver, cardiovascular disorders such as cardiac arrest,
CC cerebrovascular disorders such as cerebral ischaemia, nervous system
CC disorders such as Alzheimer's disease, infections caused by bacteria,
CC viruses and fungi, alcohol disorders such as corneal infection, endocrine
CC disorders such as premature labour and infertility, gastrointestinal
CC disorders such as Crohn's disease, renal disorders such as
CC glomerulonephritis and respiratory disorders such as asthma. The
CC polypeptides can also be used to aid wound healing, to prevent skin aging
CC due to sunburn, to maintain organs before transplantation, to regenerate
CC tissues and in chemotaxis. The polypeptides can also be used as a food
CC additive or preservative to increase or decrease storage capabilities.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published.pol_sequences.

Query Match	23.8%;	Score 912;	DB 22;	Length 176;
Best Local Similarity	99.4%;	Pred. No. 1.7e-80;		
Matches 175; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0

QY	541	AOISVALEFIDINIAAFIIIVFSGSMFYSVHOSATATEIRNOYKEMILAKKEPFI	600
Db	1	eqivsvafflignlaeffilvfygsmfysvngaicatetnqykhemilakrfive	60
QY	601	TDALCWIRIFVYKFSLSLQVEIRGTITSWYIFLPIINSALNPILYITLRPREKEMHRT	660
Db	61	fdalcwldpifvckfislqlvelpqlitswviflplinsalnpilyltctpikemihrt	120
QY	661	WYNBOKRSMDSKGOITYAPSFIVWEMPVLOEMPPELMPDLTYPCEMSLSQST	716
Db	121	wynyqirksmdsqgktyapsfiwemvmploemppeilmkpoltypcemsllsqst	176

RESULT 12

ID AU18666 standard; Protein; 176 AA.

AC AAU1.8666;

DT 21-NOV-2001 (first entry)

DE Renal and cardiovascular-associated protein, Seq ID 105.

KM	Human antiinflammatory; neuroprotective; immunomodulator; vulnereary;
KM	cardiovascular; cyostatic; nephrotoxic; antanaemic; nephritis;
KM	immunosuppressive; kidney disorder; renal failure; hypertension;
KM	cardiovascular disorder; myocardial infarction; blood disorder; anaemia;
KM	blood coagulation disorder; electrolyte imbalance disorder; cancer;
KM	hyponaemia; hyperkalaemia; neoplastic disorder; nephroma;
KM	autoimmune disease; inflammatory disease; reproductive system disorder;
KM	endocrine disorder; neural activity; neurological disorder;
KM	wound healing; respiratory disorder.

OS Homo sapiens.

PN W0200155328-A2.

PD 02-AUG-20

PF 17-JAN-2001; 2001WO-US01359.

PR	31-JAN-2000;	2000US-0179065;
PR	04-FEB-2000;	2000US-0180028;
PR	24-FEB-2000;	2000US-0184664;
PR	02-MAR-2000;	2000US-0186570;
PR	16-MAR-2000;	2000US-0189574;
PR	17-MAR-2000;	2000US-0190076;
PR	18-APR-2000;	2000US-0198123;
PR	19-MAY-2000;	2000US-0205515;
PR	07-JUN-2000;	2000US-0209467;
PR	28-JUN-2000;	2000US-0214886;
PR	30-JUN-2000;	2000US-0215135;
PR	07-JUL-2000;	2000US-0216647;

PR	07-JUN-2000	2000US-0216680
PR	11-JUL-2000	2000US-0217487
PR	11-JUL-2000	2000US-0217496
PR	14-JUL-2000	2000US-0218290
PR	26-JUL-2000	2000US-0220963
PR	26-JUL-2000	2000US-0220964
PR	11-AUG-2000	2000US-0224518
PR	14-AUG-2000	2000US-0224519
PR	14-AUG-2000	2000US-0225415
PR	14-AUG-2000	2000US-0225447
PR	14-AUG-2000	2000US-0225457
PR	14-AUG-2000	2000US-0225758
PR	14-AUG-2000	2000US-0225759
PR	18-AUG-2000	2000US-0226681
PR	22-AUG-2000	2000US-0226688
PR	22-AUG-2000	2000US-0227182
PR	23-AUG-2000	2000US-0227189
PR	30-AUG-2000	2000US-0229824
PR	01-SEP-2000	2000US-0229827
PR	01-SEP-2000	2000US-0229943
PR	01-SEP-2000	2000US-0229944
PR	01-SEP-2000	2000US-0229945
PR	05-SEP-2000	2000US-0229959
PR	05-SEP-2000	2000US-0229953
PR	06-SEP-2000	2000US-0230437
PR	06-SEP-2000	2000US-0230438
PR	06-SEP-2000	2000US-0231142
PR	08-SEP-2000	2000US-0231143
PR	08-SEP-2000	2000US-0231144
PR	08-SEP-2000	2000US-0231143
PR	08-SEP-2000	2000US-0231144
PR	08-SEP-2000	2000US-0232080
PR	08-SEP-2000	2000US-0232081
PR	12-SEP-2000	2000US-0231568
PR	14-SEP-2000	2000US-0232397
PR	14-SEP-2000	2000US-0232398
PR	14-SEP-2000	2000US-0233065
PR	21-SEP-2000	2000US-0234323
PR	21-SEP-2000	2000US-0234374
PR	25-SEP-2000	2000US-0234697
PR	25-SEP-2000	2000US-0234698
PR	26-SEP-2000	2000US-0235844
PR	27-SEP-2000	2000US-0235834
PR	27-SEP-2000	2000US-0235837
PR	29-SEP-2000	2000US-0236526
PR	29-SEP-2000	2000US-0236527
PR	29-SEP-2000	2000US-0236567
PR	29-SEP-2000	2000US-0236568
PR	29-SEP-2000	2000US-0236569
PR	29-SEP-2000	2000US-0236570
PR	02-OCT-2000	2000US-0236602
PR	02-OCT-2000	2000US-0237037
PR	02-OCT-2000	2000US-0237038
PR	02-OCT-2000	2000US-0237039
PR	02-OCT-2000	2000US-0237040
PR	13-OCT-2000	2000US-0239935
PR	13-OCT-2000	2000US-0239937
PR	20-OCT-2000	2000US-0240960
PR	20-OCT-2000	2000US-0241121
PR	20-OCT-2000	2000US-0241785
PR	20-OCT-2000	2000US-0241786
PR	20-OCT-2000	2000US-0241787
PR	20-OCT-2000	2000US-0241808

[illegible]

PR	08-NOV-2000	2000US-0246477	
PR	08-NOV-2000	2000US-0246478	
PR	08-NOV-2000	2000US-0246523	
PR	08-NOV-2000	2000US-0246524	
PR	08-NOV-2000	2000US-0246525	
PR	08-NOV-2000	2000US-0246526	
PR	08-NOV-2000	2000US-0246527	
PR	08-NOV-2000	2000US-0246528	
PR	08-NOV-2000	2000US-0246532	
PR	08-NOV-2000	2000US-0246538	
PR	08-NOV-2000	2000US-0246609	
PR	08-NOV-2000	2000US-0246610	
PR	08-NOV-2000	2000US-0246611	
PR	08-NOV-2000	2000US-0246613	
PR	17-NOV-2000	2000US-0249207	
PR	17-NOV-2000	2000US-0249208	
PR	17-NOV-2000	2000US-0249209	
PR	17-NOV-2000	2000US-0249210	
PR	17-NOV-2000	2000US-0249211	
PR	17-NOV-2000	2000US-0249212	
PR	17-NOV-2000	2000US-0249213	
PR	17-NOV-2000	2000US-0249214	
PR	17-NOV-2000	2000US-0249215	
PR	17-NOV-2000	2000US-0249216	
PR	17-NOV-2000	2000US-0249217	
PR	17-NOV-2000	2000US-0249218	
PR	17-NOV-2000	2000US-0249244	
PR	17-NOV-2000	2000US-0249245	
PR	17-NOV-2000	2000US-0249264	
PR	17-NOV-2000	2000US-0249265	
PR	17-NOV-2000	2000US-0249297	
PR	17-NOV-2000	2000US-0249289	
PR	17-NOV-2000	2000US-0249300	
PR	01-DEC-2000	2000US-0250160	
PR	01-DEC-2000	2000US-0250391	
PR	05-DEC-2000	2000US-0251030	
PR	05-DEC-2000	2000US-0251988	
PR	05-DEC-2000	2000US-0256719	
PR	06-DEC-2000	2000US-0251479	
PR	08-DEC-2000	2000US-0251856	
PR	08-DEC-2000	2000US-0251868	
PR	08-DEC-2000	2000US-0251869	
PR	08-DEC-2000	2000US-0251889	
PR	08-DEC-2000	2000US-0251990	
PR	11-DEC-2000	2000US-0254097	
PR	05-JAN-2001	2001US-0259678	
XX			
PA	(HUMA-)	HUMAN GENOME SCI INC.	
XX			
PI	Rosen CA, Barash SC, Ruben SM;		
XX			
DR	WPI: 2001-465558/50.		
DR	N-PSDB; AAS34845.		
XX			
PT	Novel polypeptides and polynucleotides useful as diagnostic reagents to		
PT	diagnose diseases or disorders associated with aberrant expression or		
PT	activity of polypeptides, and for treating cancers, rheumatoid		
PT	arthritis		
XX			
PS	Claim 11; SEQ ID NO 373; 687pp; English.		
XX			
CC	The present invention relates to the isolation of novel human neoplastic		
CC	disease associated polypeptides, and cDNA (AAS34767-AAS35050) and DNA		
CC	sequences encoding for these polypeptides. The sequences of the		
CC	invention are useful in the diagnosis, treatment, prevention and/or		
CC	prognosis of disorders involving neoplastic disease such as		
CC	hyperproliferative disorders (e.g. leukemia, bone cancer, bladder		
CC	cancer, brain stem glioma, adult liver cancer, childhood cerebellar		
CC	astrocytoma, or Hodgkin's lymphoma). The sequences of the invention may		
CC	also be useful for treating other disorders such as neural disorders,		
CC	immune system disorders, muscular disorders, reproductive disorders,		
CC	gastrointestinal disorders, pulmonary disorders, cardiovascular disorders,		
CC	and renal disorders. The polynucleotide sequences of the invention are		
CC	also useful in gene therapy. AAU21568-AAU21851 represent the novel human		

CC neoplastic disease associated polypeptides of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 176 AA:

Query Match 23.8%; Score 912; DB 22; Length 176;
Best Local Similarity 99.4%; Pred. No. 1.7e-80;
Matches 175; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 541 AQLYSAIFGILNLAFLIVSGSMFVSQSAITAEIRNOVKEMIAKREFFIVF 600
Db 1 AQLYSAIFGILNLAFLIVSGSMFVSQSAITAEIRNOVKEMIAKREFFIVF 60
QY 601 TDALCWFIFVFEVFLSLQVEIPGTTISWVIFILPINSALNPILVLTTRPEKEMIHRE 660
Db 61 tdaLcwiflvfEVLslqVeIpGttiswVifilPinsalNPilVlttrPeKemiHrE 120
QY 661 WYNYRQRKSDSGQKTYAPSFIVWEKMPLEKMPLEKMPLEKMPLEKMPLEKMPLEKMPLE 716
Db 121 WYNYRQRKSDSGQKTYAPSFIVWEKMPLEKMPLEKMPLEKMPLEKMPLEKMPLEKMPLE 176

RESULT 14

ID ABB04105 standard; Protein; 172 AA.

XX ABB04105;

DT 08-JAN-2002 (first entry)

DE Human musculoskeletal system related polypeptide SEQ ID NO 2052.

KW Cytostatic; immunosuppressive; nocotropic; neuroprotective; antiviral;
KW anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; antilucer;
KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein;
KW musculoskeletal system.

XX Homo sapiens.

PN WO200155367-A1.

PD 02-AUG-2001.

PF 17-JAN-2001; 2001WO-US01338.

XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234224.
PR 25-SEP-2000; 2000US-0234297.
PR 25-SEP-2000; 2000US-0234598.
PR 25-SEP-2000; 2000US-0234599.
PR 26-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 27-SEP-2000; 2000US-0235837.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240360.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.

PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246600.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249246.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251088.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX
XX WPI; 2001-451937/48.
XX
XX N-PSDB; AAL35687.
XX
XX Isolated polypeptide for treating, preventing and/or prognosing
PT disorders related to the musculoskeletal system including
PT musculoskeletal cancers and also for testing and detection e.g.
PT diagnosis -
XX
XX
PS Claim 11; SEQ ID NO 2052; 781pp + Sequence Listing; English.

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 172 AA;
Query Match 18.7%; Score 717.5; DB 22; Length 172;
Best Local Similarity 88.8%; Pred. No. 1,5e-61;
Matches 142; Conservative 3; Mismatches 14; Indels 1; Gaps 1;
QY 326 RMFRPLMNLSHIEKFKFOYGYAPHRSCCKPNTDGLISLENLASTIHFQFVWVSAVTC 385
|||||
Db 1 RMFRPLMNLSHIEKFKFOYGYAPHRSCCKPNTDGLISLENLASTIHFQFVWVSAVTC 60
|||||
QY 386 FGNIFVYICMRPIYRSNKLYAMSIIISLCADCMGIVLFVIGGFDLKFRGEYKHAQLW- 444
|||||
Db 61 FGNIFVYICMRPIYRSNKLYAMSIIISLCADCMGIVLFVIGGFDLKFRGEYKHAQLW 120
|||||
QY 445 MESTHCOLVGSALILSTEVSVLLITPLEKTCIYPPR 484
| : ||||| : : ||||| : : |||||
Db 121 GEYSIXAKKDXLALISTEVSKLIVTXLTLEKNICIDHPFR 160
RESUME 15
AAB41526
ID AAB41526 standard; Protein; 140 AA.
XX
XX AAB41526;
AC
XX
XX 08-FEB-2001 (first entry)
DT
XX
XX Human ORFX ORF1290 polypeptide sequence SEQ ID NO:2580.
DE
XX
XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnary; antiposrotic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antineoplastic;
KW antiviral; antibacterial; antifungal; antineoplastic; antihypertensive;
KW antineoplastic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive.
XX
XX Homo sapiens.
OS
XX
XX WO200058473-A2.
XX
XX 05-OCT-2000.
PD
XX
XX 31-MAR-2000; 2000WO-US08621.
PF
XX
XX 31-MAR-1999; 99US-0127607.
XX
XX 02-APR-1999; 99US-0127636.
PR
XX 05-APR-1999; 99US-0127728.
PR
XX 30-MAR-2000; 2000US-0540763.
XX
XX (CURA-) CURAGEN CORP.
PA
XX
XX Shinkets RA, Leach M;
XX
XX WPI; 2000-602362/57.
XX
XX N-PSDB; AAC75735.
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX
PS Claim 11; Page 1840-1841; 5507pp; English.

CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnertary;
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC antihypoid; and antianaemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.

XX Sequence 140 AA;

Query Match 17.0%; Score 650; DB 21; Length 140;

Best Local Similarity 99.2%; Pred. No. 4e-55;

Matches 127; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 552 INIAFIIVFSYGSMEFVSQSAITATEIRNOVKKEMILAKRFFIVFTDALCWIPIFV 611
|||
DB 8 iniaafliivfsgsmfysvghgsaitateirngvkkemilakrffivfdalcwipifv 67
|||
OY 612 VKFLSLQVEIPGTISWVVIITLPINSALNPILYLTTRPFKEMIHREWYNYRQRKSM 671
|||
DB 68 vkflslqveipgtlswvviitlpinsalnpllylcltrpfkemihrfwyrqrksmd 127
|||
OY 672 SKGQKTYA 679
|||
DB 128 skgqktea 135
|||

Search completed: September 5, 2002, 08:25:16
Job time: 10676 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2002, 06:27:55 : Search time 66.25 Seconds
(without alignments)
266.193 Million cell updates/sec

Title: US-09-647-067-8
Perfect score: 3834
Sequence: 1 MTRGSVFFILIRGKYFHSK.....FTYPCMSLISOSTRLNSYS 722

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCNUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfilest.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	483.5	12.6	696	US-07-757-342D-4	Sequence 4, Appl
2	481.5	12.6	692	US-07-757-342D-6	Sequence 6, Appl
3	478	12.5	695	US-08-487-886-2	Sequence 2, Appl
4	478	12.5	695	US-08-487-885-2	Sequence 2, Appl
5	470.5	12.3	700	US-07-757-342D-3	Sequence 3, Appl
6	469	12.2	674	US-07-757-342D-10	Sequence 10, Appl
7	469	12.2	699	US-07-757-342D-2	Sequence 2, Appl
8	449.5	11.7	611	US-07-757-342D-8	Sequence 8, Appl
9	449.5	11.7	636	US-07-757-342D-7	Sequence 7, Appl
10	441	11.5	795	US-07-741-453A-55	Sequence 55, Appl
11	439	11.5	764	US-07-741-453A-59	Sequence 59, Appl
12	438	11.4	764	US-07-741-453A-61	Sequence 61, Appl
13	437	11.4	764	US-07-757-342D-5	Sequence 5, Appl
14	430	11.2	764	US-07-741-453A-54	Sequence 54, Appl
15	426.5	11.1	792	US-07-741-453A-56	Sequence 56, Appl
16	395	10.3	764	US-07-741-453A-60	Sequence 60, Appl
17	354.5	9.2	644	US-08-866-757-2	Sequence 2, Appl
18	354.5	9.2	644	US-09-153-593-2	Sequence 2, Appl
19	317.5	8.3	336	US-08-118-270-54	Sequence 54, Appl
20	317.5	8.3	336	PCT-US93-08528-54	Sequence 54, Appl
21	312	8.1	764	US-07-741-453A-29	Sequence 29, Appl
22	302	7.9	603	US-08-190-802A-50	Sequence 50, Appl
23	302	7.9	603	US-08-477-346-50	Sequence 50, Appl
24	302	7.9	603	US-08-473-089-50	Sequence 50, Appl
25	296	7.7	605	US-08-190-802A-49	Sequence 49, Appl
26	296	7.7	605	US-08-477-346-49	Sequence 49, Appl
27	296	7.7	605	US-08-473-089-49	Sequence 49, Appl

28	294.5	7.7	332	1	US-08-118-270-53	Sequence 53, Appl
29	294.5	7.7	332	5	PCT-US93-08528-53	Sequence 53, Appl
30	293.5	7.7	605	4	US-09-063-950-5	Sequence 5, Appl
31	289	7.5	327	1	US-08-118-270-55	Sequence 55, Appl
32	289	7.5	327	1	PCT-US93-08528-55	Sequence 55, Appl
33	286.5	7.5	1480	3	US-09-191-647-7	Sequence 7, Appl
34	286.5	7.5	1480	4	US-09-540-245A-7	Sequence 7, Appl
35	286.5	7.5	1480	4	US-09-540-153-7	Sequence 7, Appl
36	286.5	7.5	1480	5	PCT-US91-09055-2	Sequence 2, Appl
37	285.5	7.4	1480	4	US-09-182-024A-5	Sequence 5, Appl
38	285	7.4	1525	3	US-09-191-647-2	Sequence 2, Appl
39	285	7.4	1525	4	US-09-540-245A-2	Sequence 2, Appl
40	285	7.4	1525	4	US-09-540-153-2	Sequence 2, Appl
41	279	7.3	1523	4	US-09-182-024A-2	Sequence 2, Appl
42	278.5	7.3	353	3	US-08-986-485-6	Sequence 6, Appl
43	273	7.1	1091	3	US-08-986-485-5	Sequence 5, Appl
44	267.5	7.0	1101	3	US-08-986-485-2	Sequence 2, Appl
45	259.5	6.8	649	4	US-09-188-930-305	Sequence 305, App

ALIGNMENTS

RESULT 1
US-07-757-342D-4
; Sequence 4, Application US/0757342D
; Patent No. 6218509
; GENERAL INFORMATION:
; APPLICANT: IGARASHI, Masao
; INVENTOR: NAKAMURA, Kazuo
; TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESS: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; CUSHMAN
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/757,342D
; FILING DATE: 10-SEP-1991
; CLASSIFICATION: <unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: BUCKLEY, Linda M.
; REGISTRATION NUMBER: 31003
; REFERENCE/DOCKET NUMBER: 41226
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)523-3400
; TELEFAX: (617)523-6440
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 696 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-07-757-342D-4
Query Match 12.6%; Score 483.5; DB 4; Length 696;
Best Local Similarity 23.3%; Pred. No. 1.2e-30;
Matches 144; Conservative 133; Mismatches 261; Indels 79; Gaps 17;
QY 87 PVSNNVTAMSLQWNLIRKLPPDCKRNVHDLQKLDL-QNNKITSISIVAFRLNSLTKLY 145

[illegible]

```

?      REGISTRATION NUMBER: 31003
?      REFERENCE/DOCKET NUMBER: 41226
?      TELECOMMUNICATION INFORMATION:
?      TELEPHONE: (617)523-3400
?      TELEFAX: (617)523-6440
?      TELE: 200291 SYR UR
?      INFORMATION FOR SEQ ID NO: 6:
?      SEQUENCE CHARACTERISTICS:
?      LENGTH: 692 amino acids
?      TYPE: amino acid
?      TOPOLOGY: linear
?      MOLECULE TYPE: protein
?      SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-07-757-342D-6

```

Query Match	12.6%	Score 481.5;	DB 4,	Length 692;
Best Local Similarity	22.7%;	Pred. No. 1.8e-30;		
Matches 169;	Conservative 140;	Mismatches 272;	Indels 165;	Gaps 28;

OY	35	GNITFCETQOLHHCNVDCGQADENCCVVVLCOCMSLPJGELDMKPFVSVP5-VSSN	93
Db	13	GTGSCCHWLOCHSN-----RVFLCDDSKV-----TEIPDLPRNA	48
OY	94	TAMSLQWNLIRKLPDPCKNYHDLQKLDLQNNKTTSISYAFRGLNSLTKLYISHNRITF	153
Db	49	IELRFEVLTKRATYRPGSFAFGDEKTEIISQNDVLE-----	85
OY	154	LKPGVEEDLHLEMLITED-NHLSRISPPFYYGNSLLILVLMNVNLTPLRDLRQLQHP	212
Db	86	IEADVFSLPKLHEIRIEKANNLXYINDEAFQNLPSLTKYLITSTGTIKHLPAVHKIQSLQ	145
OY	213	RLHWLDELGN-HIHRLNLTFTSCS-NLTVLYMRKNKINHLENENTFAPLOKDELDELG-S	269
Db	146	KV-LTLDIODNNIHIVANSPMGLSPESVILIMLSKNGEELHNAFNGTQ-LDELNLSDN	203
OY	270	NKIENLPLIRKDLKELSQNLSTNPKIOQANOQFDVLYVKLSLSEIGIE-----	319
Db	204	NNLEELPNDVFOGASGPVLTIDISRTKYHSLPHGLEMLKILRARSYRLKLPMLDKFVT	263
OY	320	-----ISNIOQGM-FRPLMNL-----	334
Db	264	LMESLTYPSHCCAFANLKRQISELHPKCNSSILRQDIDMTQIGDQVSLIDEPSTYK	323
OY	335	-SHYFKKFOY--GQYAPHVSCRPNTDGISLENLASTIORFYVWVWVSATVCFGNIFV	391
Db	324	GSDMNYNEFDYDLCNEVYDV-TGCPKPKDAENPCEDJNGYINLRYLIMFISILATGTTVY	362
OY	392	ICMRPYIRSENKLYAMSII-SLCCADOLMOIGLYFVJGDFDKRGEYNRKHQALMESTH	449
Db	383	LIV--LTFQYKLLVPRPLMCNLAFADLQIGIYVLLIASVDIHRSQYHNVAVIDMQGAC	440
OY	450	COLVGSIAISTEVSULLTFLTEKTYICITYPPRCYARPGCR-----TTVLILMYTG	504
Db	441	CDAGGFVFVASELSVYTLTAILERHHNTIYHAOL-----ECKQYLRHNASVWYLGWTF	496
OY	505	FIVAFIPLSNKEFFKNYGTGOCFPLHSDTESIGAOISVAFGLGILNLAFLIIVFSY	564
Db	497	FAALFPL-----FQISSYMKVSLPM-----DIDISPLQLYMAL-LVLNALAFVVICGY	548
OY	565	GSMFEYSVQSAITVTEIRNQVKKEMILAKRFFEVFVFDALCWMJLPIVFKFLSLQVEIPG	624
Db	549	THIYLVANPIYVS-----SSOTKIAKRAATLITFDPLCMARISPALSASLKYPLI-	601
OY	625	TITSWVYIFIL--PINSALNFIYTLTTRPRKE-----MIHRF-----W	661
Db	602	TYSKAKILLVLYFYPINSCAMPETLAIFTKNBRDFILLSKFGCYEMQAOIYRTETSSAT	661
OY	662	YNYRQKSMDSKCOKTYAPSPITWEM	687
Db	662	HNFHARKSSCSAPRV-TNSVYVPL	686

```

1 RESULT 3
2
3 Sequence 2, Application US/08487886
4 Patent No. 5744448
5
6 GENERAL INFORMATION:
7 APPLICANT: Kelton, Christie Ann
8 APPLICANT: Schweickhardt, Rene Lynn
9 APPLICANT: Cheng, Shirley Yui Yen
10 APPLICANT: Nugent, No. 5744448een Patrice
11 TITLE OF INVENTION: Human Follicle Stimulating
12 NUMBER OF INVENTION: Hormone Receptor
13 NUMBER OF SEQUENCES: 2
14
15 CORRESPONDENCE ADDRESS:
16 ADDRESSEE: Stephan P. Williams,
17 ADDRESSEE: Ares-Serono, Inc.
18 STREET: Exchange Place, 37th floor
19 CITY: Boston
20 STATE: MA
21 COUNTRY: USA
22 ZIP: 02109
23
24 COMPUTER READABLE FORM:
25 MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density
26 COMPUTER: IBM PS/2, model 55 SX
27 OPERATING SYSTEM: MS-DOS version 4.0
28 SOFTWARE: VAX/VMS Massll via Kermit to IBM MS-DOS
29
30 CURRENT APPLICATION DATA:
31 APPLICATION NUMBER: US/08/487,886
32 FILING DATE: 07-JUN-1995
33 CLASSIFICATION: 514
34 PRIOR APPLICATION DATA:
35 APPLICATION NUMBER: 07/670,085
36 FILING DATE: 15-MAR-1991
37
38 ATTORNEY/AGENT INFORMATION:
39 NAME: Williams, Stephan P.
40 REGISTRATION NUMBER: 28546
41 REFERENCE/DOCKET NUMBER: US/252
42 TELECOMMUNICATION INFORMATION:
43 TELEPHONE: (617) 723-1300
44 TELEFAX: (617) 723-8923
45
46 INFORMATION FOR SEQ ID NO: 2:
47 SEQUENCE CHARACTERISTICS:
48 LENGTH: 695
49 TYPE: Amino acid
50 TOPOLOGY: Linear
51 MOLECULE TYPE: protein
52
53 FEATURE:
54 NAME/KEY: signal sequence
55 LOCATION: -17 to -1
56 IDENTIFICATION METHOD: hydrophobic
57
58 FEATURE:
59 NAME/KEY: putative amino-terminal extracellular domain
60 LOCATION: 1 to 349
61 IDENTIFICATION METHOD: similarity with other
62 IDENTIFICATION METHOD: dimeric glycoprotein receptor extracellular
63 IDENTIFICATION METHOD: domains, hydrophilic
64
65 FEATURE:
66 NAME/KEY: transmembrane domain
67 LOCATION: 350 to 613
68 IDENTIFICATION METHOD: similarity to other G
69 IDENTIFICATION METHOD: protein-coupled receptor transmembrane domains
70
71 FEATURE:
72 NAME/KEY: putative transmembrane region I
73 LOCATION: 350 to 370
74 IDENTIFICATION METHOD: similarity to other G
75 IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions
76 IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
77
78 FEATURE:
79 NAME/KEY: putative transmembrane region II
80 LOCATION: 382 to 404
81 IDENTIFICATION METHOD: similarity to other G
82 IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions
83 IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
84
85 FEATURE:

```

```

NAME/KEY: putative transmembrane region III
LOCATION: 427 to 448
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region IV
LOCATION: 469 to 491
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region V
LOCATION: 512 to 533
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region VI
LOCATION: 557 to 580
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region VII
LOCATION: 592 to 613
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative carboxy-terminal intracellular
NAME/KEY: domain
LOCATION: 614 to 678

```

[illegible]

[illegible][illegible]

Db 204 NNLEPLNDVHGASGVYIIDSTRIRHSLSPLSGLENKKIKARSTYVKKLPLEKLYA 263
QY 320 -----ISNIQRM--FRPLMNS----- 335
Db 264 LMEASLYPSHCCAFAMWRROISELHPICNKSILROEVDYWTQRGORSIAEDNESSYS 323
QY 336 ---HIYKKPKFY--CGYAPHRSCKPNTDGISSLENLSTIQQVFWVWVAVVCFQWIF 390
Db 324 RQDMTTEFEYDLCNEVVD--TCSPPDAPNCPEDIMGVILVLMFISILAITGNI 382
QY 391 VICRPIRSENKLYAMSII--SLCADCLMGITLFEVIGFDLKFGRGYNKHAOLMEST 448
Db 383 VLVI--LTSQYKLTVPFLMCNLAFLDLGIGITLLILASVDHTRKSNHYAYDMQGA 440
QY 449 HQVLGSLAISTEVSLLTFLLEKICIVYPPRCVRGKCR---TIVLLIMIT 503
Db 441 GCDAGFEYVASELSYTLTALTRMHTTHAMQL---DCKVOLRHASVVMWGWIF 496
QY 504 GEIYAFPLSKKEFKNYGTNGVCPHSHEDTESIGNQIYSVAIFGINLAFFIIVFS 563
Db 497 AFPAALPFI---FGISYMKVSTCLPM--DIDPLSQLVMSL--LVNLVLAFLVIGCC 548
QY 564 YGSMFYVHOSATATETERNQVKEMILAKRPFIVPTDALCWIPIEVVKFLSLQVEIP 623
Db 549 YIHILYLRNPNIVSS-----SSOTRIAKRMALIFTDFICMAPISFEAISASLKVPDI 602
QY 624 GTITSWVIFIL--PINSALNPILYTLTRPKE 654
Db 603 -TVSKAKILLVLFHPINSCANPFLYAITKMR 634

RESULT 5
US-07-757-342D-3
; Sequence 3, Application US/07757342D
; Patent No. 6218509
; GENERAL INFORMATION:
; APPLICANT: IGARASHI, Masao
; MINEGISHI, Takashi
; NAKAMURA, Kazuto
; TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; CUSHMAN
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/757,342D
; FILING DATE: 10-Sep-1991
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: BUCKLEY, Linda M.
; REGISTRATION NUMBER: 31003
; REFERENCE/DOCKET NUMBER: 41226
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)523-3400
; TELEFAX: (617)523-6440
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 700 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-07-757-342D-3

Query Match 12.3%; Score 470.5; DB 4; Length 700;
Best Local Similarity 22.8%; Pred. No. 1.3e-29;
Matches 149; Conservative 135; Mismatches 233; Indels 137; Gaps 23;

QY 87 PVSNSNTAMSLQNNLRKLPPDCFKNHDQKIDL-QNNKITSISIAFGLNSLTKLY 145
Db 49 PGPRAGLARLSLTLYLPVVPISQAFRGVINEVYKIEISQSDSLERENAFVNLNLSLL 108
QY 146 LSHNR-TIFELKPGVEEDHRLLEMLIEDN-----HLSRISPPTEFGNSLILVIMNV 198
Db 109 IONKNNLIYIEPGAFYNLPRIKYSICTGTGRTLPDYTKISSFE---NFILFCNULH 164
QY 199 LTRLPDKPLCOHMRPLMWLDLGNHINLRNLJFISCSNLTVLVWRKKNIHLENTPAP 258
Db 165 ITTIPGNAFQGMNDSVYLKLYGNQFEVQSHAF---NGTTLISLE---LKENIY-- 213
QY 259 LQKIDE-----LDLGSNKIENPLPIFDKLSQLNSYNPIQIQANQPFYL 307
Db 214 LEKMHSGAFQGATGPSLIDISTTKLQALPS-----HGLSEIQT-----L 252
QY 308 VKLSLSLEGIEISNIQRMFRPLM-----NLSHYFKKKFQ--- 343
Db 253 IALSYSLSKTLF---SKEKFTSLVATLTYPSHOCARNLPKKQONFSFIFENFSKQC 308
QY 344 -----YCGYAPHRSCKPNTDGISSLENLSTIQQ 373
Db 309 ESTVRKADNETLYSALFEENELSGWDYDYGFC--SPKTLQCAPEDAPNCPEDIMGVAF 366
QY 374 RFWVWVSAVTCFGNIFPVCIRPYRSENKLYAMSIISLCADCLMGITLFEVIGFDLKF 433
Db 367 RVLWLNILIIILFGLNLYLVLSRYKLYVPRFLMCNLSFADFCMGITLLILASVDSOT 426
QY 434 RGEYNKHAOLMESTHCOLVGSLLAISTEVSLLTFLLEKICIVYPPRCVRGKCR- 492
Db 427 KGOYNNHAIIDMOTGSCGCAAGFTYVASELSYTLTFLTRMHTIYAVQLODKRLRH 486
QY 493 TIVVILITITGFIYAFPLSKKEFKNYGTNGVCPHSHEDTESIGNQIYSVAIFGI 552
Db 487 AIPIMGLMFLSTLATPMLVG---ISNYMKVS-ICLPM--DVESTLSOVIISIL-L 538
QY 553 NLAFFIIVSYGSMFYVHOSATATETERNQVKEMILAKRPFIVPTDALCWIPIVY 612
Db 539 NVAPVVICACIRIRYFAVQNPDELAP-----NDOTIAKRMALIFTDFICMAPISFF 592
QY 613 KFLSLQVEIPGTTISWVIFIL--PINSALNPILYTLTRPKE---MIHRF 660
Db 593 AISAFAKVPFI--TYVNSKILLVLFYVNSCANPFLYAITFAFQDFILLLSRF 645

RESULT 6
US-07-757-342D-10
; Sequence 10, Application US/07757342D
; Patent No. 6218509
; GENERAL INFORMATION:
; APPLICANT: IGARASHI, Masao
; MINEGISHI, Takashi
; NAKAMURA, Kazuto
; TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; CUSHMAN
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk

TELEX: 200291 STRE UR
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 636 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 7:
 US-07-757-342D-7

Query Match 11.7%; Score 449.5; DB 4; Length 636;
 Best Local Similarity 22.5%; Pred. No. 5,7e-28;
 Matches 134; Conservative 134; Mismatches 259; Indels 69; Gaps 16;

QY 67 CQCQSLPGLDMMKPTSPVSSNTYAMSLQNLIRKLPDQFKNYHDLKDL-QNN 125
 DB 34 CMCVPGDALR-----CGPAGLTRLSLAVPKVLPISQAFRLNEVYKIEISQID 84
 QY 126 KITSISIVAFRGSLSLKLYLSNR-ITFLKPGVEEDLHRLLEMLIEDNHLRSISPT-- 182
 DB 85 SLERIEANAFDNLNLSEILIQNTKRLYIEGAFINLPGLKSLICNTGIRKRPDVTKV 144
 QY 183 FYGLNSLILVLMANNVTRLPDKPLCOHPRLLHLDLGNHINLRNLFITSCNLTFLV 242
 DB 145 FSSRSNFIIEICDMLHTTTPGNAPQGMNDSYTLKLYGNFEEVQSHAF-NGTTLSLE 203
 QY 243 MRKN-KINHNLNTEFAPLQKDELDELGSNKTLENPLIFDKELSQLNTSYNIOKIA 301
 DB 204 LKEVNHLEKHNKNGAFKRGATGKTONFSHSISENFKQCESTYKRVSKTLYSSMLASEL 263
 QY 302 NQDFYLVKLSLSLEGIEISNIQRMFRPLMNLSHYFKFQYCGYAPHVASCCKPNTDGI 361
 DB 264 SGMDY-----EXGFC--LPKPRCAPEPDAF 287
 QY 362 SLENLILASTIIOGFVYVWVSAVTCFQNIPIYICAMPYIRSEKRLYAMSIISLCCADLMGI 421
 DB 288 NPCEDINGYDFLRVLIWLIIMGMVLEVLITSRYKLYPRFLMCLNLSFDFCMGL 347
 QY 422 YLFIYGFDFKFRGEYKHAQIMESTHCOLVGSILISTEVSLTLPTLEKYICIV 481
 DB 348 YLLILASVDQTKQYNNHAIIDMOTGSGCSAGFTYFASLSLYTTLITLLEMHITTY 407
 QY 482 PFCVPRGKCR-TITVLLILWITGIVAFIPLSNKEFFKNYGTNGVCFPLHSDTESIG 540
 DB 408 AIHLDQRLRLHAILIMLGWLFSSLLIAMLPLVG--VSNNMKVS-ICFPV--DVEITL 460
 QY 541 AQLYSVAFIIGINLAARITIVESGSMFYGHOSAITATELRNVKREMLAKREFFIVF 600
 DB 461 SOYVITLITL-LNVVAFITICACIKIYFVRNPELMAT-----NDOTKIAKKMALILF 513
 QY 601 TDALCWPIFEVVFSLQVEIRPTITSWVIFTL--PINSALNPILYTLTTRPK 654
 DB 514 TDFTCMAPIEPFASIAAFKVPPLI-TVINSKVLVLFPINSCANPFIATFTKIQ 568

RESULT 10
 US-07-741-453A-55
 Sequence 55; Application us/07741453A
 Patent No. 6228597
 GENERAL INFORMATION:
 APPLICANT: PARMENTIER, MARC
 APPLICANT: LIBERT, FREDERIC
 APPLICANT: DUMONT, JACQUES
 APPLICANT: VASSART, GILBERT
 TITLE OF INVENTION: POLYPEPTIDES HAVING THYROTROPIN-RECEPTOR
 TITLE OF INVENTION: ACTIVITY, NUCLEIC ACID SEQUENCES CODING FOR SUCH RECEPTORS
 NUMBER OF SEQUENCES: 62
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: CUSHMAN DARBY & CUSHMAN
 STREET: 1615 L STREET, N.W.
 CITY: WASHINGTON, D.C.

COUNTRY: U.S.A.
 ZIP: 20036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/741,453A
 FILING DATE: 19911015
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: KOKULIS, PAUL N.
 REGISTRATION NUMBER: 16773
 REFERENCE/DOCKET NUMBER: 91913/1107/US/ST
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 861-3000
 TELEFAX: (202) 822-0944
 TELEX: 6714627 CUSH
 INFORMATION FOR SEQ ID NO: 55:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 795 amino acids
 TYPE: AMINO ACID
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-07-741-453A-55

Query Match 11.5%; Score 441; DB 4; Length 795;
 Best Local Similarity 22.2%; Pred. No. 3,6e-27;
 Matches 157; Conservative 131; Mismatches 230; Indels 190; Gaps 29;

QY 87 PSVSNVTAMSLQNLIRKLPDQFKNYHDLKDL-QNNKIT-----SISIVAFGLN 139
 DB 61 PGPAGIARLSLYLPVYKVPISQAFRLNEVYKIEISQSLERATHCGRIEANAFLNLL 120
 QY 140 SLKLYLSNR-ITFLKPGVEEDLHRLLEMLIEDN-----HLSRISPPPYGL 186
 DB 121 NLSELIQNTKRLYIEGAFINLPRLKLYSLICNTGIRATHCGRPDVTKISSSEF--- 177
 QY 187 NSLILVLMANNVTRLPDKPLCOHPRLLHLDLGN-----HINLRNLFITSCNLTIV 240
 DB 178 -NFIIEICDMLHTTTPGNAPQGMNDSYTLKLYGNFEEVQSHAF--NGTT 232
 QY 241 LVMRKNKINHNLNTEFAPLQKDEL-----IDLGSNKTLENPLIFDKELSQL 289
 DB 233 LISIE-----LKENYI--LEKMSGAFQATGSIIDISTKIQ-----ARATHC 275
 QY 290 NLSYNPIQIQANQFDYLVKLSLSLEGIEISNIQRMFRPLM----- 332
 DB 276 GLPSHGIESIQT-----LIALSSYSIKTLP-----SKKEFTSLVAVTLTYPSSHCAFRNLP 326
 QY 333 -----NLSHYFKKQ----- 343
 DB 327 RATHCGKKEQNFSEIFENFSKQCESTVRKADNETIYSALEENELSGMDRATHCGDYG 386
 QY 344 ---YCGYAPHVRSCKPNTDGISSLENLILASIIORVWVSAVTCFQNIPIYICMRP----- 396
 DB 387 RATHCGSPFTIQCAPPDPAFNCEDIMGAFRLVLIWLIIMLAIRGNLTVLVRATHCG 446
 QY 397 YIRSENKLYAMSI--SLCCADLMGIYLFVIGFDLFRGEYKHAQIMESTHCOLVG 454
 DB 447 LILSRKLYVPRFLMCLSPADFCMGVYLLILASVDSQTKQYNNHAIIDM-RATHCGQGTG 505
 QY 455 S-----LAILTEVSVLLTLTLEKYICIVP-----FRCVPRPKRTIYVLL 499
 DB 506 SGCGAGFTFVAFSELISVYTLTTLERMTIYAVQOLDKRLRNRATHC-GIPMLG 564
 QY 500 IWTGIVAFIPLSNKEFFKNYGTNGVCFPLHSDTESIGAQIYSAIFL-----GINL 554
 DB 565 GMLPSTLIATMPLVG--ISNNMKVS-ICLPM--DVESTLSOVYILISILIRATHCGLVN 617


```

; REGISTRATION NUMBER: 16773
; REFERENCE/DOCKET NUMBER: 91913/1107/US/ST
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 792 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-741-453A-56

```

```

Query Match      11.1%; Score 426.5; DB 4; Length 792;
Best Local Similarity 22.0%; Pred. No. 5.2e-26;
Matches 157; Conservative 127; Mismatches 229; Indels 201; Gaps 30;

```

```

QY 87 PSVSNVTAMSLQNLIRKLPPDCFKNYHDLOKDLQNNKIT-----SISYAFRGLN 139
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 57 PGRAGLSRLSLTLPYIKVIPSOAFRGRLNEVKIEISQSDSLPIGHGCKIEANAFDNL 116
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 140 SLKRLYLSNHR-ITFLKPGVEEDLHRLWL----- 168
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 117 NLSEILQNTKNIYIEPGAFTNLPLRLKYLSCNTGIRKPIGHGCLPDVTKIFSSEFNI 176
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 169 --IIEEDNLSRISPTPYGUNSLLVLMNNVLRLPDKPLCQHMRLHMDLEGN----- 222
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 177 LEICDNLHITVPANAFQGMN-----NEST-----LKLNGFEP 212
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 223 --HINLNLTFTSCSNLTFLVMRKKNIN--HL--NENTFAPLOKLDLDSNKE-- 273
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 213 IGHGELIQSHAF-----NGTLLISLELKENHAKMNDAFRGARGPSILDISSTKLQAPI 268
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 274 ---NLPLIFKDKEL--SOLNLSYNPQIKIQANQPDYLVKLSLEGIETISNIQRM 327
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 269 GHGCLPFGLESIQTLTATSSYSIKLPSREKFTNLDAITLTPSHCC-----A 317
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 328 FRPL-----MNLSHYFKKFO-----YC 345
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 318 FRNLPIGHGCTKEQNSFSIFKNSKQCESTARPNNEFLYSAIFAESELSMDPIGHG 377
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 346 GY-----APHVRSCKPNTDGISLENLASIIQRFVWVWSAVTCGNIFFV-- 392
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 378 GYDYGPIGHGFCSPKTLQCAPEDAFNPCEIDINGYDLKVLINILALMGVTVLFA 437
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 393 ---CMRPYLRSENKLYAMSII--SLCCADCLMGILYLVIGGFDLKRGEYKHAQLM-- 444
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 438 PIGHC--GLTSHYKLVVPRFLMCLNSFADFQCMGLYLLILASVDAQTKGOYNNHAIDMPI 495
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 445 ---MESTHQQLVGSALILSTEVSULLFTLEKTYCIYPP-----FRCYRPGCKR 492
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 496 GHGCGQTGNGSVAGFTVFSELSVYTLVYTLERWHTIYVAIOLDOKLRLRHAPIGHG- 554
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 493 TITVLILMITGFLVAFIPLSNKEFFKNYGTNGVCPPLHSEPTESGAQIYSAIFL-- 550
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 555 GIPMLGWLPSFLIAMLPLVG-----VSSYMKSYICLPM--DYETLSQYIITLILPI 607
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 551 ---GINLAFIILVSYSGSMFYSVHOSATITATEIRNOVKKEMILAKRFFITFDALCWT 607
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 608 GHGGLANVAIFIIICACYIKIYFAVQNPDELMAF-----NKDTKIAKKMAVLIFDFCMA 661
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 608 PI---FYVKFLSL-LOYEIP-GITISWVYFIL--PINSALNPLIYLTITRPER 654
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 662 PIGHGCPISFAISALKVPLITVNSKVLVLVLEYPVNSCANPPLYAIFTKAFR 715
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2002, 08:16:10 ; Search time 61.54 Seconds
(without alignments)
1127.340 Million cell updates/sec

Title: US-09-647-067-8

Perfect score: 3834
Sequence: 1 MMSGVFFYIILFKYFSHG.....FTYPCMLISQSTRLNSTNS 722

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	832	21.7	113.5	2	S40241	G protein-coupled
2	501	13.1	665	2	JC1493	foliitropin recept
3	498	13.0	695	2	I45896	foliicle stimulat
4	490	12.8	907	2	JG0193	G protein-coupled
5	484.5	12.6	694	2	JC2237	foliitropin recept
6	482.5	12.6	694	2	JC4301	foliitropin recept
7	481.5	12.6	692	2	A34548	foliitropin recept
8	481.5	12.6	696	2	A41344	foliitropin recept
9	481	12.5	695	1	JN0898	foliitropin-chorlogon
10	478	12.5	695	1	QRHUT	foliitropin recept
11	476.5	12.4	907	2	JE0176	orphan G protein-c
12	475.5	12.4	696	2	JC7361	foliitropin recept
13	471.5	12.3	700	2	A42395	lutropin receptor
14	470.5	12.3	700	2	I77463	lutelinizing hormon
15	470.5	12.3	700	2	A49744	lutropin-chorlogon
16	469	12.2	699	1	QRHUT	G protein-coupled
17	451.5	11.8	925	2	JC2033	thyrotropin recept
18	447.5	11.7	764	2	I48882	thyrotropin recept
19	443	11.6	793	2	JC7390	thyroid stimulat
20	440.5	11.5	764	2	JC5643	thyroid stimulat
21	436	11.4	764	1	QRHUT	thyrotropin recept
22	434	11.3	764	2	A35956	thyrotropin recept
23	430	11.2	764	2	A40077	thyrotropin recept
24	418	10.9	814	2	JC7389	thyroid stimulat
25	370	9.7	889	2	T20123	hypothetical prote
26	302	7.9	603	2	JC1282	insulin-like growt
27	296	7.7	603	2	A41915	insulin-like growt
28	294	7.7	603	2	JC6128	insulin-like growt
29	293.5	7.7	605	2	JC5239	insulin-like growt

30	286.5	7.5	1469	2	B36665	slit protein 2 pre
31	286.5	7.5	1480	2	A36665	slit protein 1 pre
32	285	7.4	1523	2	T13953	MEG's protein - ra
33	283.5	7.4	1531	2	T42218	slit-1 protein hom
34	283	7.4	1389	2	T13852	gene wheeler prote
35	281	7.3	961	2	T23395	hypothetical prote
36	280	7.3	1385	2	T13887	tir protein - fru1
37	275.5	7.2	536	2	A34801	lysine carboxypept
38	273.5	7.1	1066	2	T15864	hypothetical prote
39	273	7.1	1091	2	A58532	glial cell membran
40	269.5	7.0	707	2	JC7763	neuronal leucine-r
41	268	7.0	361	2	A53860	chondroadherin pre
42	266.5	7.0	610	2	T23836	hypothetical prote
43	261	6.8	594	2	T23841	hypothetical prote
44	260	6.8	1051	2	T13174	gpi30 protein - fr
45	253.5	6.6	1134	1	A29944	chaoptin precursor

ALIGNMENTS

```
RESULT 1
S40241
G protein-coupled receptor - great pond snail
C:Species: Lymnaea stagnalis (great pond snail)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 15-Jun-2001
C:Accession: S40241
R:Tennessen, C.P.; Kesteren, E.R.; Planta, R.J.; Cox, K.; Burke, J.F.; Heerikhuizen, H.;
submitted to the EMBL Data Library, June 1993
A:Description: A G protein-coupled receptor with LDL-binding motifs suggests a role f
A:Reference number: S40241
A:Accession: S40241
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1115 <TEN>
A:Cross-references: EMBL:Z23104; NID:9438128; PID:9438129
C:Superfamily: great pond snail LDL receptor-related G protein-coupled receptor; LDL
C:Keywords: G protein-coupled receptor; transmembrane protein
F:38-77/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F:79-113/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F:118-153/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F:158-194/Domain: LDL receptor ligand-binding repeat homology <LDL4>
F:195-230/Domain: LDL receptor ligand-binding repeat homology <LDL5>
F:233-267/Domain: LDL receptor ligand-binding repeat homology <LDL6>
F:274-316/Domain: LDL receptor ligand-binding repeat homology <LDL7>
F:322-361/Domain: LDL receptor ligand-binding repeat homology <LDL8>
F:367-401/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:406-440/Domain: LDL receptor ligand-binding repeat homology <LDL10>
F:446-483/Domain: LDL receptor ligand-binding repeat homology <LDL11>
F:488-523/Domain: LDL receptor ligand-binding repeat homology <LDL12>
F:584-607/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F:608-631/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F:632-655/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F:656-679/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F:704-727/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F:774-797/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>

Query Match 21.7%; Score 832; DB 2; Length 1115;
Best Local Similarity 30.0%; Pred. No. 7.7e-48;
Matches 205; Conservative 114; Mismatches 240; Indels 124; Gaps 13;
```

```

Db 27 CSLGFFPGGNTKCLPOLLHCNGVDDC-GNADADENCYV-----VICQC---MSLPG 75
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 488 CLEQFRCRK-SFQINQTKVCDGTVDCLOGMWDENCRNYWCPSHOAICQCGVMDTCGQ 546
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 76 ELDMWKPTSVSVSSANTSLQNLIRKLPPQCFKRYHD-LQKLDQNNKITSISIA 134
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 547 KLEKEN---PYQMEEDSKIMIGDNLNLTSTFSATYKVVYLDLSRNLHLEIPYIS 602
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 135 FRGINSLTLYLSNRRITFLKPGVEEDLHRLDWLIIEDNHLRSISPPTYGLNSLILVL 194
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 603 FQNNMKLTHLNLADNNITSLKNG----- 625
```

```

QY 195 MNNVLRRLDPKPLCOHMPRLHMLDLEGNHILRLNLFTISCSNLTVLVARKKINHLNEN 254
Db 626 -----SLGLSLWKLQHLNKGKIEFTIED 649
QY 255 TRAPLOKDELDELGSKKIENLPPLIFKOLKELSQLNLSTNPLOKIOANOPDYIVKLSLS 314
Db 650 TESSMHLTVLTLNSQRLNHLHYKKNMGKLOITVNLISRQINSIDNGAFNNLANVLR 709
QY 315 LSGEISNTIOQMRPRLMLSHIYEFKKFOYCGYAPHRVSCRPNTDGISLENLSTIQR 374
Db 710 LSGNVIKIDIGQYKFMGLPRLVLEKIDSYRCCIAPEGVKCSPKQDFFSCEDLMSHVLR 769
QY 375 VFWVWVAATVCGNIFVICMRPIRSEKNTLYAMSIISLCCADCMIGIYLFVIGDFLKR 434
Db 770 VSLWVLGVALVGNFVIMRWVDFRGKVHSLFLTNLAIQDLMGVYLLIINTADTYR 829
QY 435 GEYNKHAQIMESTHCQVLSALISTEVSVLLTFLTEKYICYPRCPVRCRCRTI 494
Db 830 GYVISHDENWKKOSGLCOFAGFVSTFSELSVLLSTITDRLCILEPRLRFLGRQAI 889
QY 495 TYVLIHWINGFIATFPLSNKEPFKNYCGNGVCPHSDPESIGAQIYSVAIFLGINL 554
Db 890 IYMSCIWLTVFLAVLPPLGFTFENFYGRSGYCALHVTDPDRPGWE-YSGVFTLLNL 948
QY 555 AAEIIVFSGSMFYSV---HOSAITATETIRNOVKREMLAKRFFPIVFDALCWIPFV 611
Db 949 LSFVLLASSYLMWF-SVAKKTRSAVRTAESKN---DNAMARMTLLVMDFCWPVITV 1003
QY 612 VWFSLQLQVEIPGTTISWVVIPLPINSALNPILTYLTRPF-----KEMIR 659
Db 1004 LGRVSLAGARADQVYMAVFLPLNSATNPVITYTLSTAPLGNVRKANRFRKSFHS 1063
QY 660 F-----WYNRORKS 669
Db 1064 FTGDTKHSYVDGCTHSTCEKKS 1086

RESULT 2
JC1493
folliculin receptor - sheep
M:Alternate names: follicle stimulating hormone receptor
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 21-Jan-2000
C:Accession: JCI493; I47080
R:Khan, H.; Yarney, T.A.; Saitram, M.R.
Biochem. Biophys. Res. Commun. 190, 888-894, 1993
A:Title: Cloning of alternately spliced mRNA transcripts coding for variants of ovine le
A:Reference number: JCI493; MUID:93176195
A:Accession: JCI493
A:Molecule type: mRNA
A:Residues: 1-695 <KHA>
A:Experimental source: testis
R:Yarney, T.A.; Saitram, M.R.; Khan, H.; Ravindranath, N.; Payne, S.; Seidah, N.G.
Mol. Cell. Endocrinol. 93, 219-226, 1993
A:Title: Molecular cloning and expression of the ovine testicular follicle stimulating h
A:Reference number: I47080; MUID:93351750
A:Accession: I47080
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-695 <YAR>
A:Cross-references: GB:L07302; NID:q165864; PIDN:AAA1525.1; PID:q165865
C:Genetics:
A:Gene: FSH-R
C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h
C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
F:71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>
F:191/Binding site: carbohydrate (Asn) (covalent) #status Predicted

Query Match 13.1%; Score 501; DB 2; Length 695;
Best Local Similarity 24.4%; Pred. No. 6.5e-26;
Matches 168; Conservative 130; Mismatches 245; Indels 146; Gaps 28;

```

```

QY 40 CLPQLIHC-NGVDDCGNQADENCVYVLCQCSMLPGLLELDMMKPFITSVPS-VSSNTAMS 97
Db 18 CHRLCHCSNG-----VFLCQDSKV-----TEPSSDLPRAVELR 52
QY 98 LQMWLRKLPPDCKFKYHDLQKLDLQNNKITSISYAFGLNSLFTLYSHNRITFLKPG 157
Db 53 FVLTKRLVPEGAFSGFGLKEITISQNDLEV-----IEAN 89
QY 158 VFEDRLRLMLIED-NHLSRISPPFTYGLNSLILVLMNNVTLRLDPKPLCOHMPRLM 216
Db 90 VESNLPKLEIRKNNNLITIDPDAFQNLPLNRYLLISNTGKHLPAVHKIQSLQKV-L 148
QY 217 LDLEGN-HIHLNRLTEFISCSNLTVLV-MRKNKINHLNENTRAPLOKDELDELGSN-KIE 273
Db 149 LDIQDNIHITVERNSFMGLSFESMIWLSKNGIOEIHCAFGTQ-LDELNLSDMSNLE 207
QY 274 NLPPLIFKD-----LKEISQMLNSTP-----IQIOANOPDYIVKLSLS-LERGIEI 320
Db 208 ELRPNDVEGASGPVILDISRTIRSLPSYGLENLKLRAKSTYHKLKPLSLEKFTVLEA 267
QY 321 S-----NIQOMFRPLMNL-----H 336
Db 268 SLTYPSHCAPAMWRQOTDLHPICNKSILRQEVDMQARGRISLADDEPSYAKGF 327
QY 337 IYFKRFQY-CGYAPHRVSCRPNTDGISLENLSTIORVFWVWVAATVCGNIFVICM 394
Db 328 MMYSEFDYDLCSEVVDV-TCSPEPDAFNPCEDIMGYDILRVLIWFTISLITAGNILLVVI 386
QY 395 RPYIRSEKNTLYAMSI--SLCCADCMIGIYLFVIGFEDLKFRGEYKHAQIMESTHCQL 452
Db 387 --LITSQYKLVPRFLMCMNFAFDLCIGYLLILASVDVHTKQYINVAIDMGTAGCDA 444
QY 453 VGSLLISTEVSVLLTFLTEKYICYPRCPVRCRCR-----PIVLIHWINGFI 507
Db 445 AGFTYFASGLSYTLFATLIERMHTITAMQL---ECKVHRHAASIMLVGWVAFAY 500
QY 508 AFIPLSNKEPFKNYGTNGVCPHSDPESIGAQIYSVAIFGINLAFTIIVESYGS 567
Db 501 ALFPI-----FGISSYMKVSGICLPM--DIDSPLSQLYVMSL-LVLVLAFFVYICGCTHI 552
QY 568 FYSVHOSAITATETIRNOVKREMLAKRFFPIVFDALCWIPVWFLSLOVEIGTIT 627
Db 553 YLTVRNPNITSS-----SSDTKIRAMAMLITTDLCNAPISFFAISLSKPYLI-TVS 605
QY 628 SWVYFIRL-PINSALNPILTYLTRPF 654
Db 606 KSKILLVLFYPINSCANPFLYAFITRNR 634

RESULT 3
I45896
follicle stimulating hormone receptor - bovine
C:Species: Bos primigenius taurus (cattle)
C>Date: 15-Oct-1996 #sequence_revision 15-Oct-1996 #text_change 21-Jan-2000
C:Accession: I45896
R:Ronde, A.; Lambert, A.; Saumande, J.; Silversides, D.W.; Lusier, J.G.
Mol. Reprod. Dev. 39, 127-135, 1994
A:Title: Structure of the bovine follicle-stimulating hormone receptor complementary
A:Reference number: I45896; MUID:95127199
A:Accession: I45896
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-695 <KOU>
A:Cross-references: GB:L22319; NID:q404671; PIDN:AAC37324.1; PID:q404672
C:Genetics:
A:Gene: FSHR
C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repea
F:71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Query Match 13.0%; Score 498; DB 2; Length 695;
Best Local Similarity 24.5%; Pred. No. 1e-25;

```


F:146-189/DomAin	leucine-rich alpha-2-glycoprotein repeat	homology <LRr5>
E:172-193/DomAin	leucine-rich alpha-2-glycoprotein repeat	homology <LRr6>
F:194-218/DomAin	leucine-rich alpha-2-glycoprotein repeat	homology <LRr7>
F:366-386/DomAin	transmembrane	#status predicted <TM1>
F:398-420/DomAin	transmembrane	#status predicted <TM2>
F:443-464/DomAin	transmembrane	#status predicted <TM3>
F:485-507/DomAin	transmembrane	#status predicted <TM4>
F:528-549/DomAin	transmembrane	#status predicted <TM5>
F:573-596/DomAin	transmembrane	#status predicted <TM6>
F:608-639/DomAin	transmembrane	#status predicted <TM7>
F:191,199,268,293/Binding site:	cabdohydrate (Asn) (covalent)	#status predicted

Query Match	12.6%;	Score 484.5;	DB 2;	Length 694;
Best Local Similarity	-24.1%;	Pred. NO. 8.2e-25;		
Matches 166;	Conservative 130;	Mismatches 246;	Indels 147;	Gaps 27;

```

QY 40 CLPOLLHNGVDDGNCNADENCCVYVLCQCCSLFGLFELDMKKPPTSPS-VSSNVYAMS L 98
Db 18 CHHRCVHCSSN-----RVFLCQ-----ESKVTLPSPDPRALFLRF 53
QY 99 QMNLIRKLPPDCFKNYHDLQKIDLPNNKITSISIVAFRGINSILKTLSHNRITFLKPGV 158
Db 54 VLTNRVAVIPKAFSGFGLKIEIISQNDVLEV-----IEANY 90

```

[illegible]

```

Db      209 LPNDVFGQASPVLLDISGTRHISLPNPGLENLKKLRARSTYNIKKRPLPLEKVALMEAN 268
Oy      314 -----SLEGTISNIIOQRMFRPLMNTSH----- 366
Db      269 LTPSHCCAFANMRQRSELTOTCNKSLIREVDMTOARGRR--VSLAEDEDESSPKGFD 366
Oy      337 IYFKKFOY--CGAPAHYRSCPKNTDGISLSENLASTIIOREVVNVVAVQCGNIFPICM 364
Db      337 MMYSEPRVDTGNEVNDV--AGSPEDVDAVMDQPIKCYITIVIVITGEGITATNTNTIVIV 366

```

[illegible][illegible]

QY	628	SWVVFIL--PINSALNDILVTTTRPEK	654
Db	605	KSRLIVLFYPIINSCANPELVAITTKNR	633
RESULT	6		
JC4301			
		foliitropin receptor - plg	
		foliitropin names: foli[c]e-stimulating hormone receptor	

C:\Accession: JC4301
C:\Date: 16-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 21-Jan-2000
C:\Species: Sus scrofa domestica (domestic pig)
C:\Accession: JC4301

R:Remy, J.J.; Lahbib-Mansais, Y.; Verle, M.; Bozon, V.; Couture, L.; Pajot, E.; Grebe Gene 163, 257-261, 1995
A:Title: The porcine follicleotropin receptor: cDNA cloning, functional expression and chr
A:Reference number: Jc4301; MUID:96011644
A:Accession: Jc4301
A:Molecule type: mRNA
A:Residues: 1-694 <REM>
A:Cross-references: GB:L31966
A:Experimental source: ovarian granulosa cells
C:Comment: This receptor belongs to the family of the G-protein coupled receptors. It
ematogenesis in male and oogenesis in female.
C:Genetics:
A:Gene: fshr
A:Map position: 3 q2.2-q2.3
C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repea
C:Keywords: G protein-coupled receptor; hormone receptor; transmembrane protein
F:1-365/Domain: follicle-stimulating hormone binding #status predicted <HOB>
F:70-94/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>
F:366-368/Domain: transmembrane #status predicted <TM1>
F:396-420/Domain: transmembrane #status predicted <TM2>
F:443-464/Domain: transmembrane #status predicted <TM3>
F:485-507/Domain: transmembrane #status predicted <TM4>
F:528-549/Domain: transmembrane #status predicted <TM5>
F:573-596/Domain: transmembrane #status predicted <TM6>
F:608-629/Domain: transmembrane #status predicted <TM7>

```

Query Match          12.6%  Score 482.5  DB 2:  Length 694;
Best Local Similarity 24.3%  Pred. No. 1.1e-24;
Matches 170; Conservative 126; Mismatches 255; Indels 149; Gaps 28;

Oy      29  LGYFGCGNTKCLPOLLHC-NGVDCCGQADNEDNCNVYLCQCMSPGLLELDMMKPFSTVSP  87
      | : | : : | | | | | | | | | | | | | | | | | | | | | | | |
Db      9  LAFLLTGS---CHRHICHSNG-----VFLLCQ-----  32

```

[illegible]

```

Oy 206 FLLCHMFFLMMHDLBEGN H L N L N L F F I S O S N D I V A M K R N N I N F A N E N E I F R F D O R N L
Db 138 HKIOSLOKCV-LDLDIODNNINHTVARNSPFGMSFMSMLTMSKNGIRBIRHCAFGTQ-LD 1955
Oy 264 ELIDGSN-KIENPLPIFKDLKELSQNLN SYNPRKIOQANOVDIVYKLSLEGI- 319
Db 196 ELNLSDNNNLEELPNDVYHGASGVILIDISTRHSLPSYGLEMLKKLRLARSTYNNLKKLP 255
Oy 320 -----INSIQORM-FRPLMNLSHI----- 337

```

[illegible]

Db 375 AITNNIETLVI--LITSQKLTVPRLMCMIAFDLCIGITLLILIASIDIHTKSQYHNVA 4322

QY 442 QUMNESTHCQVLSIALISTFVSLLLTFLTLEKXICIVYPERVRKGR-----11V 4366

Db 433 INMOTGACDAGGEFTVPASLSTYTLTATLTLEEMHRTIAMOL---OCKVOYRAAST 4488

QY 497 LILIMGEIYAFPLSKKEPFKYVYCTNGCPELHSDREDSIGAOIYSAIFLGIMAA 5566

Db 488 MYETATPEFTVITPDI-----PCTSSVMWVSGIDPM---DIDSLIOLYVST--LVNVA 5408

0y 557 FIIVESYSGMFYSVHQSAITATEIRNOVKKEMILAKREFEFTDALCWPPIPVKKLS 616
|::| | :::| | :: : ||| ::||| | ::|
20 405 MUIQIETIE I VUUEFI KQISSINIVSICPEM DDDSEDOYLIVCE LUNIVER 574

```
Db      541 FVYIGCGVIHYLYVRNNIMSS-----SSDTKAKRMAMLIFFDLCPNPISFALSA    594
Oy      617 LLOVEIPGTTTSWVFETL--PINSALNPIILYTTPRPFK    654
       | : | : | : ||| | | | | : | : | : 
Db      595 SLKVPLI-TVSEKLTLVLFPINSCANPFVAJFTKNFR    633

RESULT   7
A34548
follictrophin receptor precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 13-Aug-1999
C:Accession: A34548; A41729
R:Sprengel, R.; Braun, T.; Nikolics, K.; Segaloff, D.L.; Seeburg, P.H.
Mol. Endocrinol. 4, 525-530, 1990
A>Title: The testicular receptor for follicle stimulating hormone: structure and function
A:Reference number: A34548; MUID:91125358
A:Accession: A34548
A:Molecule type: mRNA
A:Residues: 1-692 <SP>
A:Cross-references: GB:I02842; NID:g204183; PIDN:AAA41175.1; PID:g204184
R:Heckert, L.L.; Daley, I.J.; Griswold, M.D.
Mol. Endocrinol. 6, 70-80, 1992
A>Title: Structural organization of the follicle-stimulating hormone receptor gene.
A:Reference number: A41729; MUID:92149579
A:Accession: A41729
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-692 <HEC>
A:Cross-references: GB:S61198; NID:g245344; PIDN:AAB2415.1; PID:g245345
A>Note: sequence inconsistent with the nucleotide translation
A>Note: Sequence extracted from NCBI backbone (NCBIN:81117, NCBIN:81119, NCBIN:81121, NCBIN:81122).
R:Davis, D.; Liu, X.; Segaloff, D.L.
Mol. Endocrinol. 9, 159-170, 1995
A>Title: Identification of the sites of N-linked glycosylation on the follicle-stimulating hormone receptor
A:Reference number: A57562; MUID:95295729
A:Contents: annotation; glycosylation sites
A:Function:
A:Description: receptor that mediates the biochemical effects of follitropin
C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h
C:Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; hormone rece
F:1-15/Domains: signal sequence #status predicted <SIG>
F:16-692/Product: follitropin receptor #status predicted <MAT>
F:16-366/Domains: extracellular hormone binding #status predicted <EHB>
F:56-70/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LR3>
F:71-95/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
F:96-120/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LR3>
F:121-145/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LR4>
F:146-169/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LR5>
F:172-193/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LR6>
F:194-218/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LR7>
F:367-387/Domains: transmembrane #status predicted <TM1>
F:398-421/Domains: transmembrane #status predicted <TM2>
F:443-465/Domains: transmembrane #status predicted <TM3>
F:486-508/Domains: transmembrane #status predicted <TM4>
F:529-550/Domains: transmembrane #status predicted <TM5>
F:571-597/Domains: transmembrane #status predicted <TM6>
F:609-650/Domains: transmembrane #status predicted <TM7>
F:191,199,293/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:534/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
F:595/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted

Query Match          12.6% Score 481.5 DB 2 Length 692;
Best Local Similarity 22.7%; Pred. No.1.3e-24;
Matches 169; Conservative 140; Mismatches 272; Indels 165; Gaps 28;

Oy      35 GNITKTCLPOLHAGNVDCGNQADENCVVVLCOCMSLPGLDLMKRFFTGPS-VSSNV    93
       | : | : | : | : | : | : | : | : | : | : | : | : 
Db      13 GTSGSCHHWLVCHSNI-----RVFLCDGSKY-----TELPIDLPANA    48
       | : | : | : | : | : | : | : | : | : | : | : | : 

Oy      94 TAMSLQMWLIKRPDPCEFKNYHDLOKDLDLNKKNTITSIVAFRGINSLTKLVLSHNRITTF    153
       | : | : | : | : | : | : | : | : | : | : | : | :
```

[illegible]

[illegible]

F:367-387/Domain	transmembrane	#status	predicted	<TM1>
F:399-421/Domain	transmembrane	#status	predicted	<TM2>
F:444-465/Domain	transmembrane	#status	predicted	<TM3>
F:486-508/Domain	transmembrane	#status	predicted	<TM4>
F:529-550/Domain	transmembrane	#status	predicted	<TM5>
F:574-597/Domain	transmembrane	#status	predicted	<TM6>
F:609-630/Domain	transmembrane	#status	predicted	<TM7>
F:191-199, 293, 318/Binding site:	carboxylate (Asn)	(covalent)	#status	predicted
F:555/Binding site:	phosphate (Thr)	(covalent)	(by protein kinase C)	#status
F:596/Binding site:	phosphate (Ser)	(covalent)	(by protein kinase C)	#status

Query Match	12.5%;	Score 481;	DB 1;	Length 655;
Best Local Similarity	24.8%;	Pred. No. 1.4e-24;		
Matches 158;	Conservative 125;	Mismatches 238;	Indels 116;	Gaps 25

OY OMNRKRLPPDCKFKYHVLQKLDLONNKITSISYAFGLSNIFLYLASHRI--TFLPKG 157
 Db 33 QESKATELPSDLPRRAIELR---FVHTKLRVYQKQAFSGFDLEKIEISQNDVLEVIAD 89
 OY 158 VFEDIHLREMLTIED-NHLSRISPPTEYGLNSLLVLVNNVLRPLRDLKPLCOHMRPH- 215
 Db 90 VFSNLPKHLHEIRIEKANNLLINPEAFONLPLRLLINSTGIKHLPD-----VHKHS 143
 OY 216 ---WLDLGN-HINLNLTNLFISCS-NLTYLVMKRKNINLNENTFAPLOKDELDIG- 268
 Db 144 FOKVILLDIQDNINHTIRNSFVGLSFESVILMLKNGIOELHNCANGTO-LDELNLSD 202
 OY 269 SNKIENTPLPFKDKELQSOLNLSYNPQIOQANOFDVLVRLKSLS-----LEGIE--- 319
 Db 203 NNNLELNDVDFHGAASGVILLIDISRTIRHSLPSYGLMKLKLRASTVNLAKPLSLEKLV 262
 OY 320 -----ISNIQRM--FRPLNLS----- 335
 Db 263 ALMEASLTYPSSHCAFAWMRRQISELHPICKNKSILRQEVDMYQTRGQRSSLAEDNESSY 322
 OY 336 ---HIEKKFQY--CGYAPHVRSCKPMTDGISLLENLASSIQRFVWVAVYCFGNI 389
 Db 323 SRGPMYTAEDPYDLCNEVDV-ITSPKPDANPCEDDILGNILRLVLMFISILAITGNI 381
 OY 390 FVICKRPYIRSEBNLYAMSI-----ISLCADCIMGILYVYIGGFDEKFRGEYNKHAOLW 444
 Db 382 IVL-----VLTLSQSKYKLTVPREFLMCNLAFADLCIGIYLLLIASVDIHTKSOYHNVAIDW 436
 OY 445 MESTHCQLVGLSALISTEVSULLTFLFLKCYICVYFRCVREPCKR-----TTVYL 499
 Db 437 QTGACGDAAGFFTYFASELSYTTLATTLERHHTTHAMOL-----DCVHVRHNASVAM 492
 OY 500 IWTGFIYAFIPLSKKEFFKNYGTNGVCPFLHSEDTSGISAGIYSAVIFGLINLAETI 559
 Db 493 GMIFAFAPALPFI---FGISSYMKVSICLPM--DIDPSLQLYVMSL-LVLNVLAFV 544
 OY 560 IYFSGSAPFYSHOSAITLRIRNOVKKEMILAKREFIYVTDALCWPIPIVAVFLSLIQ 619
 Db 545 ICGCTHLYTLVRNDNIVSS-----SSDRIARMAMLIFTDFELCAPISFPAISASLK 598
 OY 620 VEIPGTISWVIFIL--PINSALNPILYTTTRPFK 654
 Db 599 VPLI-TVSKAKILLVLFYINSCANPFLYAITTKNR 634
 RESULT 10
 ORHDFT
 follicleotropin receptor precursor - human
 N:Alternate names: follicle stimulating hormone receptor (FSHR)
 N:Contains: follicleotropin receptor precursor long splice form; follicleotropin receptor pre
 C:Species: Homo sapiens (man)
 C:Date: 30-Sep-1991 #sequence_revision 06-Sep-1996 #text_change 22-Jun-1999
 C:Accession: I57661; I56448; PC1147; S30560; I57672; JN0122
 R:Gromoll, J.; Dankbar, B.; Gudermann, T.
 Mol. Cell. Endocrinol. 102, 93-102, 1994
 .Title: Characterization of the 5' flanking region of the human follicle-stimulating
 .Reference number: I57661; M01D:95011044

A:Accession: 157661
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-51 <GR>
A:Cross-references: GB:S73199; NID:g6685036; PIDN:AA832071.1; PID:g6685037
R:Gromoll, J.; Ried, T.; Holtgreve-Grez, H.; Nieschlag, E.; Gudermann, T.
J. Mol. Endocrinol. 12, 265-271, 1994
A:Title: Localization of the human FSH receptor to chromosome 2 p21 using a genomic probe
A:Reference number: 156448; MUID:9500244
A:Accession: 156448
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 286-695 <GR>
A:Cross-references: GB:S73526; NID:g668069; PIDN:AA832225.1; PID:g668070
R:Gromoll, J.; Gudermann, T.; Nieschlag, E.
Biochem. Biophys. Res. Commun. 188, 1077-1083, 1992
A:Title: Molecular cloning of a truncated isoform of the human follicle stimulating hormone
A:Reference number: FC1147; MUID:93075197
A:Accession: FC1147
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-223,286-294, 'P',296-342 <GR3>
A:Cross-references: EMBL:X68044; NID:g31473; PIDN:CAA48179.1; PID:g31474
A:Experimental source: testis
R:Gromoll, J.
submitted to the EMBL Data Library, August 1992
A:Accession number: S30560
A:Molecule type: mRNA
A:Residues: 1-12, 'R',14-223,286-294, 'P',296-342 <GR4>
A:Cross-references: EMBL:X68044; NID:g31473; PIDN:CAA48179.1; PID:g31474
R:Keltou, C.A.; Cheng, S.V.; Nugent, N.P.; Schweickhardt, R.L.; Rosenthal, J.L.; Overton
Mol. Cell. Endocrinol. 89, 141-151, 1992
A:Title: The cloning of the human follicle stimulating hormone receptor and its expression
A:Reference number: 157672; MUID:93246012
A:Accession: 157672
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-679, 'N',681-695 <KEL>
A:Cross-references: GB:S59900; NID:g300072; PIDN:AA826480.1; PID:g300073
R:Minegishi, T.; Nakamura, K.; Takakura, Y.; Ibuki, Y.; Igataishi, M.
Biochem. Biophys. Res. Commun. 175, 1125-1130, 1991
A:Title: Cloning and sequencing of human FSH receptor cDNA.
A:Reference number: JN0122; MUID:91222171
A:Accession: JN0122
A:Molecule type: mRNA
A:Residues: 1-111, 'T',113-196, 'AV',199-306, 'R',308-695 <MTN>
A:Cross-references: EMBL:M65085; NID:g182770; PIDN:AA52477.1; PID:g182771
A:Genetics:
A:Gene: GDB:FSHR
A:Cross-references: GDB:127510; OMIM:136435
A:Map position: 2p21-2p16
A:introns: 223/3
A>Note: the exact position of the intron cannot be determined from the experimental data
C:Function:
A:Description: receptor that mediates the biochemical effects of follitropin
C:Superfamily: glycoprotein hormone receptor: leucine-rich alpha-2-glycoprotein repeat h
C:Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; hormone receptor; intron
F:1-695/Product: follitropin receptor precursor, long splice form #status predicted <SPI>
F:1-223,286-695/Product: follitropin receptor precursor, short splice form #status predicted <SG>
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-695/Product: follitropin receptor #status predicted <MAT>
F:16-366/Domain: extracellular hormone binding #status predicted <EMB>
F:56-70/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F:71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F:96-120/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F:121-145/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F:146-169/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F:172-193/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F:194-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
F:367-387/Domain: transmembrane #status predicted <TM1>
F:398-421/Domain: transmembrane #status predicted <TM2>
F:444-465/Domain: transmembrane #status predicted <TM3>

Query Match 12.5%, Score 478; DB 1; Length 695;
Best Local Similarity 23.1%; Pred. No. 2,26-24;
Matches 160; Conservative 129; Mismatches 250; Indels 154; Gaps 26;

F:486-508/Domain: transmembrane #status predicted <TM>
F:529-530/Domain: transmembrane #Status predicted <TM>
F:574-597/Domain: transmembrane #status predicted <TM>
F:609-630/Domain: transmembrane #status predicted <TM>
F:191,199,293,318/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:555/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
F:596/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted

QY 40 CLPQLHCNGVDDCGNQADENCVVYLCCQMSLPGLLEDMKPTSVS-VSSNVTAMSL 98
| : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : :
Db 18 CHNRHCHCSN-----RVFLCQ-----BSKYELPSDIPRNAIELPRF 53
QY 99 QMNLIRKLPPDFCKRYHDLOKLDLONNKRTISITYAFGLNSLRKLYSHNRITFLKPGV 158
| : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : :
Db 54 VLTKRIVYOKGAFSGFSFGDEKEIISQNDLEY-----IEADV 90
QY 159 PEDHLRLEMILED-NHLSRISPPTYGLNSLLILVMNNVLTPLDPKPLCOHMRLL-- 215
| : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : :
Db 91 FSNLRLKEIRIERKANNNLTYINPEAFOMLPLOYLLISNTGIKHLPD-----YHKIHSL 144
QY 216 ---WIDLEGN-HIHNLRLTFSCS-NLTVMRKNTINHNMENFAPILOKDELIDG-S 269
| : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : :
Db 145 OKVLDDIDONNIHTIERNSFYGLSFESVIIMLNKGIOEIHNCANFTGO-LDELINLSDN 203
QY 270 NKIENLPRLFKDLELSQLNSYNPIOKIQANDPYLVTKLSLEIG----- 319
Db 204 NMLELPMDVFHGAGPYILDISRIRHSLSYGLENKKIKRARSTYMALKPLTEKLVA 263
QY 320 -----ISNIQQRK-FRPLMNLS----- 335
Db 264 LMEASLYPSHOCAPANNRRQISELHPICNKSILRQEDVMTQTGRQSIAEDNESSYS 323
QY 336 --HYFFKKFOY-CGYAPHVRSCKPNFDGISSENLASTIORVFWVAVTCFGNIF 390
| : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : :
Db 324 RGFDMTVEFDLDLCNEVDV-TGSPKPDARNPCGDINGVINLARLVIFSIATIGNII 382
QY 391 VICMRPYRSENKLYAMSII-STCCADCLMGIVLFYVGEGDLKFRGEYNKHAOLMEST 448
| : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : :
Db 383 VLVI-LTTTSQYKLVPRPFMLCMNAFADLCIGIYLLLASVDIHRKSOYHNAYAIMDCGA 440
QY 449 HQQLGSALISTEVSLLFPLETKYICYYPRCVRPGCR-----TIIVLIKIT 503
| : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : :
Db 441 GCDAAAGFTVPASELSVYTLTAITLERMHTITHANQL----DCKQYLAHASVAMWGIF 496
QY 504 GTIVAFLPSNKEPFKNYYGTNGVCPLHSEDTESIGAQIYSVALFIGLINLAFTIIFS 563
| : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : :
Db 497 AFAAALFPFI---FGISSMKVYSICLP--DIDBPLSQLYVMSL-LVLNALAFVVICG 548
QY 564 YGSMFYSVHQSAITTELIRNOVKREMIILAKREFEFVFDALCWPIPVFKELSLQVEIP 623
| : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : :
Db 549 YHHITLVTRNPITVS-----SSPTRIAKKRAMLIFDFLOMAPISEFPAISASLKVPLI 602
QY 624 GTITSMVVIFIL-PINSALNPILYTLTPRPK 654
| : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : :
Db 603 -TVSKAKITLVLPDRINSCANPELYAITTKNR 634

RESULT 11
JEO176
orphan G protein-coupled receptor precursor - human
C.Species: Homo sapiens (man)
C.Date: 03-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 21-Jul-2000
C.Accession: JEO176
R.McDonald, T.; Wang, R.; Bailey, W.; Xie, G.; Chen, F.; Caskey, C.T.; Liu, Q.
Biochem. Biophys. Res. Commun. 247, 266-270, 1998
Affile: Identification and Cloning of an orphan G protein-coupled receptor of the gl
A.Reference number: JEO176; MUID:98308104

A:Molecule type: mRNA
A:Residues: 1-907 <MCD>
A:Cross-references: GB:AF062006; NID:g3366801; PIDN:ACC28019.1; PID:g3366802
C:Comment: This protein is a receptor for a novel class of glycoprotein ligands.
C:Genetics:
A:Gene: HG38
A:Map position: 12q22-23
F:1-21/Domain: signal sequence #status predicted <SIG>
F:562-583/Domain: transmembrane #status predicted <TM1>
F:594-616/Domain: transmembrane #status predicted <TM2>
F:639-660/Domain: transmembrane #status predicted <TM3>
F:681-701/Domain: transmembrane #status predicted <TM4>
F:725-744/Domain: transmembrane #status predicted <TM5>
F:768-791/Domain: transmembrane #status predicted <TM6>
F:803-824/Domain: transmembrane #status predicted <TM7>

Query Match 12.4% Score 476.5; DB 2: Length 907;
Best Local Similarity 23.0% Pred. No. 3.9e-24;
Matches 195; Conservative 134; Mismatches 316; Indels 203; Gaps 28;

```
QY 28 SLGYPCCGNIT-----KCLPOLHCGVDCGQADENCVVVLCQCMSLPGLELDWMP 82
DB 101 ALTYIPKGFETGLYSKVL--MLQNNQLHVPPEA-----LQYLRSLQSLRDANI 150
QY 83 FTSVPSVSSNVTAMSLQW--NLIRRLPPDCFRKNYHDQKLDLQNNKITSISYAFRGIN 139
DB 151 SYVPPSCFGLSHLRMLMDNALTEIPQAFRSLAQAMTLALNKIHIDPYAGNLS 210
QY 140 STKLYLSNNRTFLKPGFEDLHRLLEMLIEDNHSRISPPFYGLNSLILVLMNNVL 199
DB 211 SLVVLHLNHNRLHSIGKCFDGLHSLLETLDLVNNINLDEF--PFAIRTLNMLKELGFHSNNI 269
QY 200 TRLPDKPLC-----OHMRLHMLDLEG--NTHINLNLNLFISC 235
DB 270 RSIPKEAFPNQSLTIHFHYDNPICQVGSFQHPBELFTLTLNGASQITPEPDLT--GT 327
QY 236 SNLTVLVLMKRNKINHLNENTFAPLQKLDLDELGSNKIENMLPPIELFKDLKELSQLNSYNP 295
DB 328 ANLESITLGAQISLPQYVNCQPLQVLDLSYNLELDPSS--FSVCQKIQKIDLRHNE 385
QY 296 IOKIQANODIYLVKLKSLSLLEGIEISINQOQMRPLMN-----L 334
DB 386 IYEIKVDYTFQQLSLKSLMLAMNKIAIHPNASTPLSLIKLDLSNLLSPFTLGHGL 445
QY 335 SHIYFK-----KPYCGY----- 347
DB 446 THUKLTGNALQSLISSENPPELKYTEMRYAQCAFGVCENAYKITSNQMNKGDNSSMD 505
QY 348 -----APHYRSCKPNTDGISSLENLTIORVFW 378
DB 506 LHKKDGAMFQADDERLDELDFLEEDLRKALHSVQCSPPGPKPECHLDGLWLRIGYW 565
QY 379 VVS--AVTQPGNITV---ICMRYISEKKIYAMSTISLCCADCLMGIFYLVIGSDLKF 433
DB 566 TIAVALTLC--NALVSTYFRSPLYSPIKLL--IGVIAAVMMLGVSSAVLAGVADFT 620
QY 434 RGEYNKHAOLMESTCOLVGLSLAISTEVSLLTFLEKTYICIFYFRVCRPKCKRT 493
DB 621 PGSFARHGAMWENGVCYHIGLSTIFASSESVFLTLALALERFSYKAKFEPTKAPSS 680
QY 494 ITVLILIM--ITGFIYAFIPLSNKEFFKNYGTNGVCFPLHSEDTESIGQIYSVALFLGI 552
DB 681 LKVIITLCLALLTMAAVPLLG-----GSKYGASPLCLPIPGEPSTMG--YVVALIL-L 732
QY 553 NLAAPFIIVFSVGSMPYSYHQSATITATEIRNOYKKEMIIAKRFFIYVFDALCWITFIYV 612
DB 733 NSLCFMMMTIATKLYCNLDKG-----DLENWDCSMV--KATALLFPLNCILNCVAPL 785
QY 613 KFLSLQVQV--IPGITSWVFIPLINSALNPILTYLTRPFKEMIIHRF-----MYNTR 665
DB 786 SFSSLIINTLPISEVYIKFILLVVPPLPACINPLLIYLFNPHKEDLVSLRKQIYVWTRSK 845
```

```
QY 666 -----QRKSND--KGQKTYAPSFIVEMXPIQEMPELMKPDLFYPCMSLI 712
DB 846 HPSLMSINSDDVKOSCDSTQALVTFSSITVDLPSSVSP-----ATP-----V 892
QY 713 SOSTRLNS 720
DB 893 TESCHLSS 900
```

RESULT 12
JC7361
folitropin receptor precursor - newt
M:Alternate names: follicle-stimulating hormone receptor
C:Species: Cynops pyrrhogaster (newt)
C:Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 31-Dec-2000
C:Accession: JC7361
R:Nakayama, Y.; Yamamoto, T.; Oba, Y.; Nagahama, Y.; Abe, S.
Biochem. Biophys. Res. Commun. 275, 121-128, 2000
A:Title: Molecular cloning, functional characterization, and gene expression of a fo
A:Reference number: JC7361
A:Contents: Testis
A:Accession: JC7361
A:Molecule type: mRNA
A:Residues: 1-696 <NAK>
A:Cross-references: DDBJ:AB005587
C:Comment: This protein, containing seven transmembrane domains and a large glycosyla
in and thyroid stimulating hormone. This receptor has a common signal transduction pa
C:Genetics:
A:Gene: fsh-r
C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repea
C:Keywords: disulfide bond; glycolysis; glycoprotein; hormone receptor; testis; trans
F:1-11/Domain: signal sequence #status predicted <SIG>
F:18-996/Product: follicle-stimulating hormone receptor #status predicted <EXT>
F:18-359/Domain: extracellular #status predicted <EXT>
F:370-389/Domain: transmembrane #status predicted <TM1>
F:402-424/Domain: transmembrane #status predicted <TM2>
F:447-468/Domain: transmembrane #status predicted <TM3>
F:489-511/Domain: transmembrane #status predicted <TM4>
F:532-553/Domain: transmembrane #status predicted <TM5>
F:577-600/Domain: transmembrane #status predicted <TM6>
F:612-633/Domain: transmembrane #status predicted <TM7>
F:465-190,198,267,292/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:445-520/Disulfide bonds: #status predicted

Query Match 12.4% Score 475.5; DB 2: Length 696;
Best Local Similarity 23.4% Pred. No. 3.3e-24;
Matches 163; Conservative 135; Mismatches 250; Indels 149; Gaps 30;

```
QY 32 FPGCNITKCLPOLHCGVDCGQADENCYVVLDCQCMSLPGLELDWMPKPTSVS 90
DB 16 FCGHPYCRCLINRYTC-----QESHVY-----QIPRDP 44
QY 91 SNVTAMSLQNNLIRKLPDPCFRKNYHDQKLDLQNNKITSISYAFRGLNSITKLYLSH 148
DB 45 KNSTELREVLTKYVYIPKAAFGSEFDEVENITISQNDVLTITRANVSHLPKLRITIEKA 104
QY 149 NRTFLKPGVFEDLHRLLEMLIEDNHSRISPPFYGLNSLILVLMNNVFLRLPDKPIC 208
DB 105 NMLVYIDPAPQNLPLSKYLLSNTGIQV--PAVSKIRS-----FHSVL----- 147
QY 209 QHMPRLHMLDLEGN--HIHULRNITPISCSNLYLV--MRNKNINHLNENTFAPLQKIDEID 266
DB 148 -----LDVQDNIMIRHIGKNFSFGLSESITIRLNKNGIEIQNHAFNGTH--LNEIN 198
QY 267 LGSN--KIENMLPPIFD-----LKEISQNLISYNP-----IOKIQANOFDLYVKLSUS 314
DB 199 LSDNQRLKLPQVFOGANGPYILDISRIIRHPLPANGIENIKKFAKRNYYIKKLPIPE 258
QY 315 --LEGIE-----ISNIQQRW--FRPLMN-----L 334
DB 259 KPAELIEANLTYPSHCCAPANERKRKSEMHPIQCNKSFQGHDAEKPEKDNLRFSNEDYL 318
```

```

OY 335 SHIYRK-----KFOY---CGAPHYRSCKNFTDGISLSEMLASIIORVWVYSA 382
Db 319 SSGYGSYSLVENGEDEFNYDYLTCNEV--HDVLCFPRKPAFNCPEIDMGNTLRVLIIMLSI 377
OY 383 VTCCGNITVICMRPIYREKNLKYAMSIIT--SLCCADCLMGTYLEVIGGFDLKEFGENVKH 440
Db 378 LAIGNITVLVI--LISQYKFYTPRFMLCMNARADLCMGITYLILLISVDTKTSQYNNH 435
OY 441 AOLMNESTHCOLVGSLSLALISTEVSVLLTLPFLTEKYICIVYFRCVRRPGKCRITTVLLI 500
Db 436 AIDMQTSGCAAGFFYVAFASBLSYTLFTVITIERMHTITVAMQLDRKVRNRHAFALMV 495
OY 501 -WITGFYIYAFPLPSKKEFFKNYGTNGVCPPLHSEDEISGAQIYSAVIFGILAAFI 559
Db 496 GWIFFYFVALIPI-----FGVSSTYKVSICLPM--DIESRSQSYIVEI--LVLWCAFLI 547
OY 560 IVFSGSMFYSHOGSATYATATEIRNOVKKEMILAKRFFIPTVTDALCWPPIVYKFLSQ 619
Db 548 ICACIGIYLYVRNPNVVS-----NSDTYKARKMALITITDPLCAPISFPAISASLK 601
OY 620 VEIPGTITSWVIFIL--PINSALNPLIYLTTPRPK 654
Db 602 IPLI-TVSKSKILLVLFYFINSKANPELYAIFTYKTR 637

RESULT 13
A42395
Luteotropin receptor - mouse
N:Alternate names: lutealizing hormone-choriogonadotropin receptor
C:Species: Mus musculus (house mouse)
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 13-Aug-1999
C:Accession: A42395
J:Gudermann, T.; Birnbaumer, M.; Birnbaumer, L.
J. Biol. Chem. 267, 4479-4488, 1992
A:Title: Evidence for dual coupling of the murine luteinizing hormone receptor to adenylyl
      none receptor expressed in L cells.
A:Reference number: A42395; MUID:92165799
A:Accession: A42395
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-700 <GDU>
A:Cross-references: GB:M81310; GB:M81318; NID:g198811; PIDN:AAA39437.1; PID:g198812
C:Note: sequence extracted from NCBI backbone (NCBIN:84064, NCBIP:84065)
C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h
C:Keywords: G protein-coupled receptor; transmembrane protein
F:54-77/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F:78-102/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F:103-127/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F:128-152/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F:153-177/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F:180-201/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F:202-226/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>

Query Match 12.3%; Score 471.5; DB 2; Length 700;
Best Local Similarity 23.8%; Pred. No. 6.2e-24;
Matches 148; Conservative 131; Mismatches 271; Indels 71; Gaps 18;

OY 87 PSVSNVTAMSLQNLNKLRLPPDCFKNYHDLQIDL--ONNKITSISIVAFGINSITKLY 145
Db 49 PPRAGLRLSLITLYPVKVIYPSQAFRGLINEVYKIEISSDSIERIANAFNLNLSEIL 108
OY 146 LSHNR-ITFLKRGVEDLRLLEWLIIEDN-----HLSRISPPFYGUNSLILVLYAMNV 198
Db 109 IQNTKNTLIYIEGATPNLRLKYLSTICNTGIRTLDPVSKISSER-----NFLTEICDLX 164
OY 199 LTRLPDKPLCOHMPRLHMLDEGNHINLRNLFTISGNSLTVLYVRKKIN---HLNEN 254
Db 165 ITTIGNAFOGMNNESTITLKLYNGFEVQSHAF---NGTTLISLEKENIYLEKMHSG 220
OY 255 TPAPLOKIDEIDGSKNKIENLP-----PLIFDKLEISQL-----NLSTY- 293
Db 221 TFQGTATGSIIDVSSTKQALPDSHGLESIQTLIATSSYLKTLPREREKFTSLVATVITP 280

```

Oy 294 -----NPICKIQAQNDYLVTKLKSLSBEGIEISNIQOQMRPLMLNSHIEFKKPOYCG 346
 Db 281 SHCAFRPLPKRQNFSSIFENFSKOCSESTVBREAN-NEIYSAIFEBENEISGMDYDIF 339
 Oy 347 YAPHRSCKPTMDGSSISLENTLAIIOREYVWVAVWCENIVICMRPYIRESENKLYA 406
 Db 340 CSPKTIQCTPPEDPAENPCEDIDMGAYFLRVLITWILINILAIKONLVLYVLLSRKRLVPR 399
 Oy 407 MSIISSCCADCMIGIYLFVIGGFDLKFERGEYVKNKAQLMESTHCOLVGSIAIISTEVSVL 466
 Db 400 FLMCNLSAFDCMGLYLLILLIASVDSQTKGYXNHAIDMOTSGCSAGGFVFAFSELSVY 459
 Oy 467 LITFLLTEKYICIVYPRFCVPRKCR-TITVLLILMITGFIYAIPLSKNEFFANNGTN 525
 Db 460 TITVTLTERRHMTTYAVOLDKRLRRAHPIMLGWMIFSTLMATLPLVG---VSTMKV 515
 Oy 526 GVCEPLHSEDESEIGAOIYVAIFGLINIAAFIIVESYSGMFSVHQSATATEIRNOV 585
 Db 516 SICLPM---DYESTLSQYIILISILL-LNAVAFVIYACAYRIYFAVQNPETLAP----- 565
 Oy 566 KKEMLARREFPIYVTDALCMIPFVYVAFSLDQVEIGTISWVIFIL--PINSALNP 643
 Db 566 NKDTIARKMALIFTDTCMAPISEFAISAFKVPIL-TVINSKVLILVLPYPVNSCANP 624
 Oy 644 ILYTLTRPKE---MIHRE 660
 Db 625 FLVAFYTKAFQORDFFLLLSRF 645
 RESULT 14
 177463
 luteinizing hormone/chorionic gonadotropin receptor - rat
 C:Species: Rattus sp. (rat)
 C:Date: 02-Aug-1996 #sequence-revision 02-Aug-1996 #text-change 21-Jan-2000
 C:Accession: J77463
 R:Aatsinki, J.T.; Pietila, E.M.; Iakkakorpi, J.T.; Rajaniemi, H.J.
 Mol. Cell. Endocrinol. 84, 127-135, 1992
 A>Title: Expression of the LH/CG receptor gene in rat ovarian tissue is regulated by
 F:153-177/Domain: leucine-rich alpha-2-glycoprotein repeat homology <RR>
 A:Reference number: 157668; MUID:92347604
 A:Accession: J77463
 A>Status: Preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-700 <RES>
 A:Cross-references: GB:S40803; NID:g252163; PIDN:AB22680.1; PID:g252164
 C:Genetics:
 A:Introns: 58/2; 82/2; 107/2; 132/2; 157/2; 183/2; 206/2; 231/2; 293/2; 320/2
 C:Superfamily: glycoprotein hormone receptor: leucine-rich alpha-2-glycoprotein repeat
 F:153-177/Domain: leucine-rich alpha-2-glycoprotein repeat homology <RR>
 Query Match 12.3%; Score 470.5; DB 2; Length 700;
 Best Local Similarity 22.8%; Pred. No.7.2e-24;
 Matches 149; Conservative 135; Mismatches 233; Indels 137; Gaps 23;
 Oy 87 PSVSSNVTAMSLQNNILIRKLPPDCFKNYHDLOKIDL-ONNKITSISIYAFGLNSLTRLY 145
 Db 49 PEPRAAGLARLSFTYPRKYVYIPQAFRGLNEVYKIEISQSDSLERIEANAFNLMLSELL 108
 Oy 146 LSHNR-TIFFKAGVEDLHRLLEWLIIEDN-----HLSRISPPPTFYGLNSLILVLMNV 198
 Db 109 IONTNLLIYIEPGATNLPRLKYLISICWTGIRTPDYKISSSEF---NPILEICDMIH 164
 Oy 199 LTRLDKPLCQMPRLHMLDEGNHIMLRNLTFTSCSNLPIVLRKKKIMHNLNENFAP 258
 Db 165 ITTIGNAPFQGMNNSVTLLKLTGNGFEFVQSAF---NGTLLISLE-----LKNIT-- 213
 Oy 259 LOKLDE-----IDLGSNKIENLPPLFKDLKELISOLNLSYNPIOKIOANOQFDYL 307
 Db 214 LEKMSGNAFQAGTGSIIIDISTKIQALPS-----HLESIQT-----L 252
 Oy 308 VKLKLSLEGIEISNIQOQMRPLM-----NLSHIEFKKPO--- 343

Db 253 IALSYSLKTLTP-----SKEKFTSLVATLTTPSHCCAFRMLPKKEQNFSEFENFSKOC 308
 QY 344 -----YCGYAPHVRSCKPNTDGISSLEMLASIIQ 373
 Db 309 ESTVRKADNETLYSAIFEENELSGWDYDGC--SPKTLCCAPPPDAFNCEDIMGYAFL 366
 QY 374 RVEVWVAVATCGNIFVICMRPYINSEKLYAMSIISLCCADCLMGITYFVIGFPLKF 433
 Db 367 RVILIMINILAIIGNLVLEFVLLTSRYKLVPRFLMCLNLSFADPCMGILYLLLASVDSQT 426
 QY 434 RGEYNKHAQIAMESTHQVLGSLAISTEVSULLFTLEKTYCIYPRPCVAPGRCR- 492
 Db 427 KGOYNNHAIIDMOWGSGGGAAGFFVFASELVTLVTLTERMTTITTYAVOLDCKLRLH 486
 QY 493 TIIVLLIWIITGFIIVAFIPLSNKEFFKNYGTGVCPLSEDESTIGAOIYSAIFLGI 552
 Db 487 AITPLMGWLFSTLIAMPLVG---ISNYMKVS-ICLPM--DVESTLSQVYILSLI-L 538
 QY 553 NLAAFIITVSYGSMFYSVHOSAITATEIRNOYKKEMLAKRFFIYFDALCWIPFVY 612
 Db 539 NVVAFVVICACYIRIYFAVONPELTAP-----NKDKRIKKAKMILIFDTFTCMAPISF 592
 QY 613 KFLSLLOVELIPGTTISWVFIIL--PINSALNPILYTLTRPREKE---MIHRF 660
 Db 593 AISAAPKVPIL-IYVTSKILVLVFPVNSCANPFLYAIPTKAFORDPFLLSRF 645

RESULT 15

A49744
 Lutropin-choriogonadotropin receptor precursor - rat
 N:Alternate names: luteinizing hormone-chorogonadotropin receptor
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 03-MAY-1994 #sequence, revision 13-Sep-1998 #text_change 13-Aug-1999
 C:Accession: A49744; A40545; A41343; A61453; A32460
 R:KOC, Y.B.; Ji, T.; Slaughter, R.G.; Ji, T.H.
 E:Enochriology 120, 2297-2308, 1991
 A:Title: Structure of the luteinizing hormone receptor gene and multiple exons of the c
 A:Reference number: A49744; MOID:91209270
 A:Accession: A49744
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1700 <KOC>
 A:Cross-references: GB:M68917
 A:Note: authors translated the codon CAA for residue 307 as Glu, AAC for residue 355 as
 R:Tsai-Morris, C.H.; Buczko, E.; Wang, W.; Xie, X.Z.; Dufau, M.L.
 J. Biol. Chem. 266, 11355-11359, 1991
 A:Title: Structural organization of the rat luteinizing hormone (LH) receptor gene.
 A:Reference number: A40545; MOID:91250455
 A:Accession: A40545
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-320 <RSA>
 A:Cross-references: GB:M63918; GB:M63919; GB:M63920; GB:M63921; GB:M63923; GE
 R:McFarland, K.C.; Sprengel, R.; Phillips, H.S.; Koehler, M.; Rosembilt, N.; Nikolics, K
 Science 245, 494-499, 1989
 A:Title: Lutropin-choriogonadotropin receptor: an unusual member of the G protein-couple
 A:Reference number: A41343; MOID:89332512
 A:Accession: A41343
 A:Molecule type: mRNA
 A:Residues: 1-700 <MCF>
 A:Cross-references: GB:M26199; NID:g205178; PIDN:AAAA1528.1; PID:g205179
 R:Dufau, M.L.; Munejishi, T.; Buczko, E.S.; Delgado, C.J.; Zhang, R.
 J. Steroid Biochem. 33, 715-720, 1989
 A:Title: Characterization and structure of ovarian and testicular LH/CG receptors.
 A:Reference number: A61453; MOID:90097014
 A:Accession: A61453
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 27-33, 'X', '35-37', 'X', '39', 'X', '41-44' <DUF>
 R:Roche, P.C.; Ryan, R.J.
 J. Biol. Chem. 264, 4636-4641, 1989
 A:Title: Purification, characterization, and amino-terminal sequence of rat ovarian rece
 A:Reference number: A32460; MOID:89174723

A:Accession: A32460
 A:Molecule type: protein
 A:Residues: 27-32, 'LX', '35-37' <ROC>
 C:Genetics: 58/2; 82/2; 132/2; 157/2; 183/2; 206/2; 231/2; 293/2; 320/2
 A:Insertions: 58/2; 82/2; 132/2; 157/2; 183/2; 206/2; 231/2; 293/2; 320/2
 C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repea
 C:Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; hormone r
 F:54-77/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR1>
 F:78-102/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
 F:103-127/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR3>
 F:128-152/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR4>
 F:153-177/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR5>
 F:180-201/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR6>
 F:202-226/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR7>
 F:368-389/Domain: transmembrane #status predicted <TM1>
 F:400-422/Domain: transmembrane #status predicted <TM2>
 F:444-466/Domain: transmembrane #status predicted <TM3>
 F:489-511/Domain: transmembrane #status predicted <TM4>
 F:530-551/Domain: transmembrane #status predicted <TM5>
 F:575-598/Domain: transmembrane #status predicted <TM6>
 F:610-631/Domain: transmembrane #status predicted <TM7>
 F:103-178, 199, 295, 303, 317/Binding site: carbohydrate (Asn) (covalent) #status predict
 F:647, 648/Binding site: palmitate (Cys) (covalent) #status predicted
 F:681/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicte

Query Match 12.3%; Score 470.5; DB 2; Length 700;
 Best Local Similarity 22.8%; Pred. No. 7.2e-24;
 Matches 149; Conservative 135; Mismatches 233; Indels 137; Gaps 23;
 QY 87 PSVSSNTAMSLQWNLRLKLPDCKFNHYDLQKLDL-QNNKITSISYAFRGINSILTKLY 145
 Db 49 PGRPAGLARLSLTLPLPVKIPSOAFRLNRYVYKIEISDSLEIRIEMARDNLINSELL 108
 QY 146 LSHNR-ITFLKPGVEFDLHLEMLIEDN-----HLSTRSPFFYNSLILLYLMNV 198
 Db 109 IQNKNLYTEPPGAFNLPRLKYLSTICNTGIRLPDVKTSSSEF---NFLIETCNLH 164
 QY 199 LTRLPDRLCOHMRPLMLDLEGNHINLRLNLFISCSNLTFLVMRKKNKINHLENTEFAP 258
 Db 165 ITTIPGNAFGMNNESTLALXGCFEEVSHNF-----NCTTILSL- -LKENLY-- 213
 QY 259 LQKIDF-----LDGSKNIENLPLRLFKDKELSQLNSIVPQIOQANOFDYL 307
 Db 214 LERKHSAGFOGAGPSILDISSTYKQLAPS-----HGLSISQF-----L 252
 QY 308 VKLKSLSLEGIEISINIOQRNFRPLM-----NLSHYFKFQ--- 343
 Db 253 IALSYSLKTLTP-----SKEKFTSLVATLTTPSHCCAFRMLPKKEQNFSEFENFSKOC 308
 QY 344 -----YCGYAPHVRSCKPNTDGISSLEMLASIIQ 373
 Db 309 ESTVRKADNETLYSAIFEENELSGWDYDGC--SPKTLCCAPPPDAFNCEDIMGYAFL 366
 QY 374 RVEVWVAVATCGNIFVICMRPYINSEKLYAMSIISLCCADCLMGITYFVIGFPLKF 433
 Db 367 RVILIMINILAIIGNLVLEFVLLTSRYKLVPRFLMCLNLSFADPCMGILYLLLASVDSQT 426
 QY 434 RGEYNKHAQIAMESTHQVLGSLAISTEVSULLFTLEKTYCIYPRPCVAPGRCR- 492
 Db 427 KGOYNNHAIIDMOWGSGGGAAGFFVFASELVTLVTLTERMTTITTYAVOLDCKLRLH 486
 QY 493 TIIVLLIWIITGFIIVAFIPLSNKEFFKNYGTGVCPLSEDESTIGAOIYSAIFLGI 552
 Db 487 AITPLMGWLFSTLIAMPLVG---ISNYMKVS-ICLPM--DVESTLSQVYILSLI-L 538
 QY 553 NLAAFIITVSYGSMFYSVHOSAITATEIRNOYKKEMLAKRFFIYFDALCWIPFVY 612
 Db 539 NVVAFVVICACYIRIYFAVONPELTAP-----NKDKRIKKAKMILIFDTFTCMAPISF 592
 QY 613 KFLSLLOVELIPGTTISWVFIIL--PINSALNPILYTLTRPREKE---MIHRF 660
 Db 593 AISAAPKVPIL-IYVTSKILVLVFPVNSCANPFLYAIPTKAFORDPFLLSRF 645

Thu Sep 5 14:39:29 2002

us-09-647-067-8.rpr

Page 11

Search completed: September 5, 2002, 09:01:18
Job time: 2708 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2002, 08:25:21 ; Search time 56.19 Seconds
(without alignments)
497.518 Million cell updates/sec

Title: US-09-647-067-8

Perfect score: 3834
Sequence: 1 MTSGSVFFYIIIFGKFSHG.....FTYPCMSLISOSTRLNSYS 722

Scoring table: BLOSUM62
Gapop 10.0 , Gapect 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	832	21.7	1115	1	GPCR_LYMST
2	505	13.2	693	1	FSHR_CHICK
3	501	13.1	695	1	FSHR_SHEEP
4	498	13.0	695	1	FSHR_BOVIN
5	488	12.7	695	1	FSHR_PIG
6	487.5	12.7	676	1	FSHR_CALJA
7	484.5	12.6	694	1	FSHR_HORSE
8	481.5	12.6	692	1	FSHR_RAT
9	481.5	12.6	696	1	FSHR_PIG
10	481.5	12.5	695	1	FSHR_MACFA
11	478	12.5	695	1	FSHR_HUMAN
12	475.5	12.4	692	1	FSHR_MOUSE
13	474.5	12.4	687	1	FSHR_EOAS
14	471.5	12.3	700	1	FSHR_MOUSE
15	470.5	12.3	700	1	FSHR_RAT
16	469	12.2	699	1	FSHR_HUMAN
17	463.5	12.1	701	1	FSHR_BOVIN
18	451.5	11.8	925	1	GLHR_ANTI
19	450	11.7	763	1	FSHR_BOVIN
20	447.5	11.7	764	1	FSHR_MOUSE
21	440.5	11.5	764	1	FSHR_SHEEP
22	437	11.4	764	1	FSHR_RAT
23	431	11.2	764	1	FSHR_HUMAN
24	430	11.2	764	1	FSHR_CANFA
25	400	10.4	538	1	FSHR_SHEEP
26	351	9.2	366	1	FSHR_MOUSE
27	302	7.9	603	1	ALS_RAT
28	296	7.7	603	1	ALS_HUMAN
29	294	7.7	603	1	ALS_MOUSE
30	293.5	7.5	905	1	TLR3_MOUSE
31	286.5	7.5	905	1	SLIT_DROME
32	286.5	7.5	1480	1	CBP8_HUMAN
33	275.5	7.2	536	1	homo sapien

34	268	7.0	361	1	CHAD_BOVIN	Q27972	bos taurus
35	260	6.8	859	1	TLR5_MOUSE	Q9J1F7	mus musculus
36	258	6.7	713	1	GAC1_HUMAN	Q753Z5	homo sapien
37	253.5	6.6	1315	1	CHAO_DROME	P12074	drosophila
38	249	6.5	331	1	PLIB_AGRBL	Q93223	agkistrodon
39	244.5	6.4	343	1	LDM_CHICK	P51890	gallus gall
40	243.5	6.4	904	1	TLR3_HUMAN	015455	homo sapien
41	241	6.3	567	1	GPV_MOUSE	008742	mus musculus
42	240	6.3	382	1	PRLP_MOUSE	P51888	homo sapien
43	235	6.1	682	1	CONN_HUMAN	001819	drosophila
44	235	6.1	782	1	CHAO_TRICA	P82963	tribolium c
45	233	6.1	966	1	Y918_HUMAN	Q94991	homo sapien

ALIGNMENTS

```
RESULT 1
GPCR_LYMST
ID GPCR_LYMST STANDARD: PRT; 1115 AA.
AC P46023;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE G-protein coupled receptor GRL101 precursor.
OS Lymnaea stagnalis (Great pond snail).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;
OC Lymnaeidae; Lymnaea.
OX NCBI_TaxID=6523;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=CNS;
RX MEDLINE=94255418; PubMed=8197140;
RA Jensen C.P., van Kesteren E.R., Planta R.J., Cox K.J.A., Burke J.F.,
  van Heerikhuizen H., Vreugdenhil E.;
RT "A G protein-coupled receptor with low density lipoprotein-binding
  RT motifs suggests a role for lipoproteins in G-linked signal
  transduction."
RT Proc. Natl. Acad. Sci. U.S.A. 91:4816-4820(1994)
RL CC
CC -1- FUNCTION: MIGHT DIRECTLY TRANSDUCE SIGNALS CARRIED BY LARGE
CC -1- EXTRACELLULAR (LIPID)PROTEIN(COMPLEXE)'S INTO NEURONAL EVENTS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN A SMALL NUMBER OF
CC -1- NEURONS WITHIN THE CENTRAL NERVOUS SYSTEM AND TO A LESSER EXTENT
CC IN THE HEART.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -1- SIMILARITY: CONTAINS 12 LDL-RECEPTOR CLASS A DOMAINS.
CC -1- SIMILARITY: CONTAINS 7 LEUCINE-RICH REPEATS (LRR).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, Z23104; CAAB0651.1; -.
CC PIR, S40241; S40241.
CC HSSP, Q07954; ICR8.
CC GCRDB, GCR_0860; -.
CC InterPro: IPR000276; GPCR_Rhodopsn.
CC InterPro: IPR002172; LDL_recept_A.
CC InterPro: IPR001611; LRR.
CC InterPro: IPR000372; LRR_Nterm.
CC InterPro: IPR003592; LRR_Out.
CC InterPro: IPR003591; LRR_Typ.
CC Pfam: PF00001; 7tm_1; 1.
CC Pfam: PF00057; LDL_recept_a; 11.
CC Pfam: PF00560; LRR_6.
CC Pfam: PF01462; LRRNT; 1.
CC PRINTS: PR00237; GPCR_RHODOPSIN.
CC PRINTS: PR00261; LDLRECEPTOR.
```


DR SMART; SM00192; LDLa; 12.
DR SMART; SM00370; LRR; 2.
DR SMART; SM00370; LRRNT; 1.
DR SMART; SM00369; LRR_TYP; 1.
DR PROSITE; PS00237; G_PROTEIN_RECIP_FL1; FALSE_NEG.
DR PROSITE; PS00262; G_PROTEIN_RECIP_FL2; 1.
DR PROSITE; PS01209; LDLa; 1; 6.
DR PROSITE; PS50068; LDLa; 2; 11.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Repeat;
leucine-rich repeat; Signal.
FT SIGNAL 1 24
FT CHAIN 25 115
FT CHAIN 25 767
FT TRANSMEM 768 788
FT TRANSMEM 789 801
FT TRANSMEM 802 822
FT TRANSMEM 823 857
FT TRANSMEM 858 878
FT TRANSMEM 879 887
FT TRANSMEM 888 908
FT TRANSMEM 909 941
FT TRANSMEM 942 962
FT TRANSMEM 963 988
FT TRANSMEM 989 1009
FT TRANSMEM 1010 1017
FT TRANSMEM 1018 1038
FT TRANSMEM 1039 1115
FT DOMAIN 32 523
FT DOMAIN 36 79
FT DOMAIN 77 115
FT DOMAIN 116 135
FT DOMAIN 156 196
FT DOMAIN 195 232
FT DOMAIN 231 269
FT DOMAIN 272 318
FT DOMAIN 320 363
FT DOMAIN 365 403
FT DOMAIN 404 442
FT DOMAIN 444 485
FT DOMAIN 486 525
FT REPEAT 557 579
FT REPEAT 582 605
FT REPEAT 607 629
FT REPEAT 630 653
FT REPEAT 655 677
FT REPEAT 678 701
FT REPEAT 703 725
FT REPEAT 725 753
FT DISULFID 38 53
FT DISULFID 46 66
FT DISULFID 60 77
FT DISULFID 79 91
FT DISULFID 86 104
FT DISULFID 98 113
FT DISULFID 118 131
FT DISULFID 138 153
FT DISULFID 158 170
FT DISULFID 165 183
FT DISULFID 177 194
FT DISULFID 202 220
FT DISULFID 214 230
FT DISULFID 233 245
FT DISULFID 240 258
FT DISULFID 252 267
FT DISULFID 267 291
FT DISULFID 274 294
FT DISULFID 282 304
FT DISULFID 298 316
FT DISULFID 322 339
FT DISULFID 334 352
FT DISULFID 346 361
FT DISULFID 367 379
FT DISULFID 374 392
FT DISULFID 386 401
POTENTIAL.
G-PROTEIN COUPLED RECEPTOR GRLL01.
EXTRACELLULAR (POTENTIAL).
1 (POTENTIAL).
CYTOPLASMIC (POTENTIAL).
2 (POTENTIAL).
EXTRACELLULAR (POTENTIAL).
3 (POTENTIAL).
CYTOPLASMIC (POTENTIAL).
4 (POTENTIAL).
EXTRACELLULAR (POTENTIAL).
5 (POTENTIAL).
CYTOPLASMIC (POTENTIAL).
6 (POTENTIAL).
EXTRACELLULAR (POTENTIAL).
7 (POTENTIAL).
CYTOPLASMIC (POTENTIAL).
12 x 40 AA APPROXIMATE TANDEM REPEATS
SIMILAR TO THE LDL-RECEPTOR CLASS A.
LDL-RECEPTOR CLASS A 1.
LDL-RECEPTOR CLASS A 2.
LDL-RECEPTOR CLASS A 3.
LDL-RECEPTOR CLASS A 4.
LDL-RECEPTOR CLASS A 5.
LDL-RECEPTOR CLASS A 6.
LDL-RECEPTOR CLASS A 7.
LDL-RECEPTOR CLASS A 8.
LDL-RECEPTOR CLASS A 9.
LDL-RECEPTOR CLASS A 10.
LDL-RECEPTOR CLASS A 11.
LDL-RECEPTOR CLASS A 12.
LRR 1.
LRR 2.
LRR 3.
LRR 4.
LRR 5.
LRR 6.
LRR 7.

FT DISULFID 406 418 BY SIMILARITY.
FT DISULFID 413 431 BY SIMILARITY.
FT DISULFID 425 440 BY SIMILARITY.
FT DISULFID 446 458 BY SIMILARITY.
FT DISULFID 453 474 BY SIMILARITY.
FT DISULFID 465 483 BY SIMILARITY.
FT DISULFID 488 500 BY SIMILARITY.
FT DISULFID 495 513 BY SIMILARITY.
FT DISULFID 507 523 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 269 269 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 318 318 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 482 482 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 502 502 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 571 571 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 618 618 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 685 685 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1115 AA; 125865 MW; E9BB01297ECE356C CRC64;
Query Match 21.7%; Score 832; DB 1; Length 1115;
Best Local Similarity 30.0%; Pred. No. 4,1e-43;
Matches 205; Conservative 114; Mismatches 240; Indels 124; Gaps 13;
QY 27 CSLGYPFCGNTKCPDLPHLCHNGVDDC-GNQAEDNCVY-----VLCQC-----MSLPL 75
DB 488 CLEGGFCRCR-SPCINOTKXCDGTVDCLGQMDENNRVCYRPGALICQCEGYMCTQG 546
QY 76 ELDMKRFPSVPSVSSNVTAMSLQWNLIRKLPDCEKRYND-LQKLDLONNKITSISIV 134
DB 547 KLEKEM-----PVQOEMEEDLSKLMIGDNLNTSTFSATYDYKYVDLDSRNLTEPIYS 602
QY 135 FRLGNSLTKLYLSHNRTRTEKPGVFEDLHLEMLIEDNLISPTPEGLMSLILVL 134
DB 603 FQNMWKTTHLNLDDNNTSLKNG----- 625
QY 195 MNVNLRLPDKPCLQHMRLHMLDEGNHINLNLFTSCSNLTVLVYMKRKNINLNN 254
DB 626 -----SLGLSNLKKOLHNGKNIETIED 649
QY 255 TPAPLQKLELDLGSNKIENLPLIFRDKELSQLNSYPIOKIOANOPDYVVKLSLS 314
DB 650 TFSMHLTVLDSNOGLRTHVYKMEFGKLQITVLNISRQINSINDGAFNNLANVRLID 709
QY 315 LBGLEISNIOQRMPRLMNSHIFYKKFOYCGYAPHYRSCKPMTDGLSSLEMLASTIOR 374
DB 710 LSGNVIKDIOQKVFMGJLRLVELKTDSPRCCLAPBGVCKSPKODESFSCEDLMSNHVLR 769
QY 375 VFVWVVSAYVCFGNIFVYICMRPYIRSEBNKLYAMSIISCCADCLMGYLVFVIGGFDLKFR 434
DB 770 VSIWVLGVILVNGFVYIFRVPDPRFGKHYSHLITRLALGDLVGLYLLIATAPIYR 829
QY 435 GEYNKAQLMESTHOLVGSIALISTEVSALLTFLEKTYICYIYFPFCVBPGRCRT 494
DB 830 GYVISHDENMKQGLQGFAGFVSTFSELSVLTSTLTDRLICLIFPLRTRILGRQAI 889
QY 495 TVLLIWLITFYIAPLPSLKEFFKNYGTNGCYCPHLASDTSISQAQISVAIFGLINL 554
DB 890 IYVSCIVWLFELAVPLGLGFSYENFGRSGVCLAHVYPRDRPWE-YSVGVFILLNL 948
QY 555 AAFITIVESGSMFYSV---HOSAITATELRNOVKKEMILAKRFFFIYVFDALCWPIPIV 611
DB 949 LSEVLIASSTLWMF-SVAKTRSAVRAESKN---DNAMARKMTLIWMTDFCCWPIIY 1003
QY 612 VKFLSLQVEIPGTTISWVVFILPINSALNPLIYLTTRPF-----KEMIR 659
DB 1004 LGFVSLAGARADQVYAMLAIVFLPLNSATNPVITYLTSLAPFLGNVRRKRANRFRKSFHS 1063
QY 660 F-----WYNYRORKS 669
DB 1064 FTGDTFRHSYVDDGTHTSYCEKKS 1086

	RESULT	2			
FSHR_CHICK	CHICK	STANDARD;	PRT:	693 AA.	
ID	FSHR_CHICK				
AC	P79763; Q90719;				
DT	01-NOV-1997 (Rel. 35, Created)				
DT	01-NOV-1997 (Rel. 35, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Follicle stimulating hormone receptor precursor (FSH-R) (Foliotropin receptor).				
GN	FSHR.				
OS	Gallus gallus (Chicken).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.				
CC	NCHI_TaxID=9031;				
OX	[1]				
RN	SEQUENCE FROM N.A.				
RP					
RC	TISSUE=Ovary;				
RX	MEDLINE=97473503; PubMed=9332357;				
RA	Wakabayashi N., Suzuki A., Hoshino H., Nishimori K., Mizuno S.; "The cDNA cloning and transient expression of a chicken gene encoding a follicle-stimulating hormone receptor."; Gene 197;121-127(1997). [2]				
RL					
RN	SEQUENCE FROM N.A.				
RP	STRAIN-WHITE LEGHORN;				
RC	MEDLINE=97057887; PubMed=8902217;				
RA	You S., Bridgman J.T., Foster D.N., Johnson A.L.; "Characterization of the chicken follicle-stimulating hormone receptor (cFSH-R) complementary deoxyribonucleic acid, and expression of cFSH-R messenger ribonucleic acid in the ovary."; Biol. Reprod. 55:1055-1062(1996). -1- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLATE CYCLASE. CC CC				
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.				
CC	-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS. FSH/LSH/TSH SUBFAMILY.				
CC	-1- SIMILARITY: CONTAINS 7 LEUCINE-RICH REPEATS (LRR).				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch). CC CC				
DR	EMBL; D87871. BAA13487.1; -;				
DR	EMBL; U51097; AAC60030.1; -;				
DR	HSSP; P23945; IXUN.				
DR	GCRDb; GCR_1077; -;				
DR	GCRDb; GCR_1137; -;				
DR	InterPro: IPR000276; GPCR_Rhodpsn.				
DR	InterPro: IPR001611; LRR.				
DR	InterPro: IPR003721; LRR_Nterm.				
DR	Pfam: PF00001; 7tm_1; 1.				
DR	Pfam: PF00560; LRR; 3.				
DR	PRINTS; PR00237; GPCRHODOPSIN.				
DR	SMART; SM00013; LRRT; 1.				
DR	PROSITE; PS00237; G_PROTEIN_RECPT_FL_1; 1.				
DR	PROSITE; PS50262; G_PROTEIN_RECPT_FL_2; 1.				
KW	G-protein coupled receptor; transmembrane; Glycoprotein; Signal; Phosphorylation; Repeat; Leucine-rich repeat.				
KW					
FT	SIGNAL	1	17	POTENTIAL.	
FT	CHAIN	18	693	FOLLICLE STIMULATING HORMONE RECEPTOR.	
FT	DOMAIN	18	366	EXTRACELLULAR (POTENTIAL).	
FT	TRANSMEM	367	387	1 (POTENTIAL)	
FT	DOMAIN	388	398	CYTOPLASMIC (POTENTIAL).	
FT	TRANSMEM	399	421	2 (POTENTIAL)	
FT	DOMAIN	422	443	EXTRACELLULAR (POTENTIAL).	

FT	TRANSMEM	444	465	3 (POTENTIAL).	
FT	DOMAIN	466	485	4 (CYTOPLASMIC (POTENTIAL).	
FT	TRANSMEM	486	508	4 (POTENTIAL).	
FT	DOMAIN	509	528	EXTRACELLULAR (POTENTIAL).	
FT	TRANSMEM	529	550	5 (POTENTIAL).	
FT	DOMAIN	551	573	CYTOPLASMIC (POTENTIAL).	
FT	TRANSMEM	574	597	6 (POTENTIAL).	
FT	DOMAIN	598	608	EXTRACELLULAR (POTENTIAL).	
FT	TRANSMEM	609	630	7 (POTENTIAL).	
FT	DOMAIN	631	693	CYTOPLASMIC (POTENTIAL).	
FT	REPEAT	45	68	LRR 1.	
FT	REPEAT	69	93	LRR 2.	
FT	REPEAT	95	118	LRR 3.	
FT	REPEAT	119	143	LRR 4.	
FT	REPEAT	168	192	LRR 5.	
FT	REPEAT	193	216	LRR 6.	
FT	REPEAT	218	240	LRR 7.	
FT	DISLIPID	442	517	BY SIMILARITY.	
FT	CARBOHYD	47	47	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	191	191	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	199	199	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	268	268	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CONFLICT	4	4	G -> D (IN REF. 2).	
FT	CONFLICT	88	88	G -> A (IN REF. 2).	
FT	CONFLICT	140	140.	K -> R (IN REF. 2).	
FT	CONFLICT	174	174	I -> T (IN REF. 2).	
FT	CONFLICT	191	191	N -> S (IN REF. 2).	
FT	CONFLICT	329	329	V -> L (IN REF. 2).	
SQ	SEQUENCE	693 AA;	78697 MW;	46F9869635A1BEC CRC64;	

Query Match	13.28	Score 505	DB 1	Length 693
Best Local Similarity	24.78	Pred. No. 1.4e-23		
Matches 1/4	Conservative 122	Mismatches 259	Indels 150	Gaps 27

QY 28 SLGYPCGNITKRLPOLHHCNGVDCGQADENCY-----VVLQCQMSLPGLLEDMKPF 83
|||-----|||
Db 2 SLG-----LTCLLILVASCSSG-----CQHHTLCBGRIFLCQELKVVOLPRD----- 43
|||-----|||

QY 84 TSVPSVSSNVTAMSLQWMLIRKLPDPCFKNYHDQLKLQNNKITSI SIYA F RGLNSLTG 143
 : : | : : : | ||| : : : |

QY 144 LYLSHNRITFLKPGVFEDIHRLLEWLIIE-D-NHLSRISPTFYGLSLLILVIMNNVLTRL 202

```

Db      81  -----DALEIEGNVFSSPLKHEIRIEREKANKLMKIDQDAFQHLPSRLYLLISNTGLSFL 1355

```

Db 136 P-----VVHKVHSFQKVALDVQDNIHRTIERNTFMGLSSSVILRLNKNGIOEIFKDHA 1893

QY 256 FAPLOKDELDELGSN-KIENLPLIFKD-----LKELSQMLISYNP-----IQKIQAQ 3033

QY 304 FDYLKVL---KSLSLEGTEIS-----NIOQRMRPLMLNSHYEKKFOYC 346

Db 249 TYKLLKKEDVKNKFRSL--IEANFTYPSHCAFTNRKTKQNTFEYPICSMSPAKQDLGEQTG 306

Db 307 KRKHRSAAEDYISHYGTFRGPEVNEFDYGLCNEVDVFCSPKDAFNPCEDIMGYNLR 366

QY 375 VFVVVWSAVTCFGNIEFLVICMRPYIRSENKLYAMSI--SLCCADCLMGIVLFEVIGGFDLK 4322

00 433 FRGEYNKHAOLMESTHCOLVGLAILSTEVSULLTFLTEKYYICIVYPRCVRPGKCR 4922

Db 425 TKSRYNNVAIDWQTGAGCNAAGFTVFASLSVYTLTVILTERWHTITYAMQLNRKVLRL 484

45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 538
 539
 540
 541
 542
 543
 544
 545
 546
 547
 548
 549
 550
 551
 552
 553
 554
 555
 556
 557
 558
 559
 560
 561
 562
 56

Query Match	13.0%;	Score 498;	DB 1;	Length 695;
Best Local Similarity	24.5%;	Pred. No. 3.7e-23;		
Matches 169;	Conservative 126;	Mismatches 248;	Indels 146;	Gaps 28;

OY		40	CLPOLHLC--NVDVDCGNQADENCNVVYLCOGSMPLSGLEIDMKRFTVS--VSSNTYAMS	97
Db		18	CHRLCHCSNG-----VELCQ-----ESKVTIEPSPDPAVELR	52
OY		98	LQWMLIRKLPPDCRNKYHDLOKLDLQNKKITTSISIAFERGLNSLTLYLSHRITTFDKPG	157
Db		53	FVLTKLVRIYPKAGSFGEDKEIKETISONDYLEV-----LEAN	89
OY		158	VFEELHREMLEIID-NHLSTISPPTYGSLILLYLMNVLTFRDPKLCOMPRLYM	216
Db		90	VFSNPKLHEIRIKANNLLTYIDPAQONLPNLXYLLISNGIHLPAVKHTOSLOKV-L	148
OY		217	LDLEGN-HIHNRILMYLFISCSNLTVLV-MRKKNINHINENTFAFLQKDELIGSN-KIE	273
Db		149	IDIQDNINIHTVERNSFMGJSFESMTYWLSKNGIQEIHNCANFTQ-LDELINTSDNSNLE	207
OY		274	NLPLPIERD-----LKELSOLNSTYNP-----IQKIQAOFDYLVUKSLS--LBCEI	322
Db		208	ELRPDVFOGASGPVILDISRTIRISLPSYGLENLKLRAKSTYRLKRLPSLEKEFVTLVEA	267
OY		321	S-----NIOQRMPRLMNTS-----H	336
Db		268	SLTTPSHOCAPAMRRQTSDLHPICNNKSILROEVDMITQARGRVSLAEDEDSYAKGFD	327
OY		337	IYEKKFOY--CGVAPHVBSCKPNTDGSISSLNLLASTIIQRYFWVWSAVTCFGNIIFYCM	394
Db		328	VMSEFPYDLCNEVADV-TCSPEPDAAFNPCEDIMGDDILRYLIWFISILTATIGNILTVI	386
OY		395	RPLYRSNKLIKAMSI-I-SLOCACMLGIYLFYVGGRDFKRGFYKNHADLMESTFCQL	452
Db		387	-LITSQKTLVPFPLMCNLAFAFUDLCIGIYILLASVDVAHTKIEYHNVAIDMOTGACDA	444
OY		453	VGSIALISTEVSALLTFLELEKTYICIVPRCYRPSKCR---TTVLILMITGYI	507
Db		445	AGEFTVAASELSYVTLTALTIERMHTITHMOUL----ECKVOLRHASIMLVGMIFAFV	500
OY		508	AFIPLSKKEPFKNYGGNGOCPRLHSDDTESIGAOIYSVALFGINLAARFIIFYSGSM	567
Db		501	ALFPI-----FGISSYMAVSILPM---DISPLSOLUYMSL-IVLVNLAFAVJICGCTHI	552
OY		568	FYSVHQSATATEIRNOVKEMILAKREFFEIVFDALCWPIEFVFKPLSLOVEIPETIT	627
Db		553	YLTVRNPNTSS-----SSDTKIAKRWAMLIFDFPLCMAPISFPAISAKVPLI-TV	605
OY		628	SWVIYFIL--PINSALNPILIYTLTRPK	654
Db		606	KSKILLVFYPINSCANPFIATPTKNFR	634
RESULT	5			
FSHR_PIG	FSHR_PIG	STANDARD;	PRT;	695 AA.
AC	p49058;_077514;			
DT	01-FEB-1996 (Rel. 33, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, last annotation update)			
DE	Follicle stimulating hormone receptor precursor (FSH-R) (Follitropin receptor).			
OS	FSHR			
OS	Sus scrofa (Pig)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.			
OX	NCBI_TaxID=9823;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Ovary;			
RA	MEDLINE=96011644; PubMed=7590277; Remy J.D., Lahib-Mansals Y., Yarle M., Bozon V., Couture L.,			

```

RA Pajot E., Griebert D., Salese R.:
RT "Pig porcine follicotropin receptor: cDNA cloning, functional
RT expression and chromosomal localization of the gene.";
RL Gene 163:257-261(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Ovary;
RA Wang Y.F., Meyer K.B., Schmidt K., Wan S.J., Degen S.J.F.,
RA la Barbera A.R.:
RT "Porcine follicle-stimulating hormone receptor.";
RL Submitted (SEP-1997) to the EMBL/Genbank/DBJ databases.
CC CC -I- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY
CC OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
CC ADENYLYLATE CYCLASE.
CC CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
CC CC -I- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC CC -I- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
CC CC FSH/LSH/TSH SUPERFAMILY.
CC CC -----
CC CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb.ch/announce/isb.ch).
CC CC or send an email to license@isb.ch).
CC CC -----
DR EMBL; L31966; AAA86933.1; -.
DR EMBL; AF025377; AAC24981.1; -.
DR HSSP; P23845; 1XUN.
DR GCRDB; GCR_1561; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000372; LRR_Nterm.
DR Pfam; PF00001; 7tm.1; 1.
DR Pfam; PF00560; LRR_4.
DR Pfam; PF01462; LRRNT_1.
DR PRINTS; PR00373; GLYCHOMONER.
DR PRINTS; PR001143; FSHRECEPTOR.
DR SMART; SM00013; LRRNT_1.
DR PROSITE; PS00237; G_PROTEIN_RECPEP_FL_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECPEP_FL_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW phosphorylation; Repeat; Leucine-rich repeat.
FT SIGNAL 1 17
FT CHAIN 18 695
FT DOMAIN 18 366
FT TRANSMEM 367 387
FT DOMAIN 388 398
FT TRANSMEM 399 421
FT DOMAIN 422 443
FT TRANSMEM 444 465
FT DOMAIN 466 485
FT TRANSMEM 486 508
FT TRANSMEM 509 528
FT TRANSMEM 529 550
FT DOMAIN 551 573
FT TRANSMEM 574 597
FT DOMAIN 598 608
FT TRANSMEM 609 630
FT DOMAIN 631 695
FT REPEAT 44 68
FT REPEAT 69 93
FT REPEAT 119 143
FT REPEAT 170 192
FT REPEAT 193 216
FT REPEAT 218 240
FT DISULFID 442 517
FT CARBOHYD 191 191
FT CARBOHYD 199 199
FT CARBOHYD 293 293
FT CONFLICT 13 13
FT CONFLICT 13 13

```


FT CARBOHYD 293 293 N-LINKED (GLCNAC...) (POTENTIAL).
S0 SEQUENCE 692 AA; 77681 MM; 267EA78C7CFD8EC6 CRC64;

Query Match 12.68; Score 481.5; DB: 1; Length 692;

Best Local Similarity 22.78; Pred. No. 3.7e-22;

Matches 169; Conservative 140; Mismatches 272; Indels 165; Gaps 28;

35 GNITKLPQLHNGVDCGNQADNDCVVLQCCMSLPGLELDMKRPFTSVPS-VSSNV 93
13 GTGSCCHMLCHCSN-----RFLCQDSKV-----TELPDLP RNA 48
94 TAMSIQMLNLRKLPDCKNTNHDLOKLDLONNKITSISIVAFRGLNSLTLYLSHNRITF 153
49 IELRVLTKLYVIRPKGSFAGGDEKIEISQNDVLEV----- 85
154 LKPYEEDLHREMLIED-NHLSRISPTFGNLSLLVLMNVNLRPLKPCLOMP 212
86 IADVFSMLPKHEIRIKANLLYINDEAFONLPSRLYLSNTGIRHLPVAKIQSLQ 145
213 RLHMLDEGN-HIHNLRNLFTIGCS-NLTVLMRKNKINHMENTFAPLOKLELDLG-S 269
146 KV-LLDIDONINIHIVANSFGLSFESVILLMSKNGIEIHNCFNGTQ-LDELINLSDN 203
270 NKIENLPILRKDKLSQLNLSYPIOKIQANQPDYLVKLSLEGIT----- 319
204 NNLELPNDVQGASGPYLDLSRTKVSLSHPNHLNKLKLRARSTYRLKPLNDKEVT 263
320 -----ISNIOQRN--FRPLMNL----- 334
264 LMEASLTTPSHCCAFANLKRQISELHPICNKSILNODIDMTQIGDQVSLIDEPSTYK 323
335 -SHYFKKQY--CGYAFHVSCKRNTDGISLENLASIQRVFVWVAVTCGNITFV 391
324 GSDMYNEFDYDLCEVVDV--TCSPPKPAFNPCEIDMGYNILRYLIMFISLAIATGNITV 382
392 ICMRYISENKLAMSTI--SLCCADLMGTYLFVIGFDEKFGENKRAQLMESTH 449
383 LVV--LTTSOYKLVYPRFLMCLNADICIGYLLLASVDIRKSOYHNVAIDMGTGAG 440
450 COLGSLAILTEVSVLLTFLTEKYICYVPRFCVARGCR-----TIVYLLIMWTG 504
441 CDAGFFVYVSELSVYTLTTLERKHTITHAMQ-----ECKVOLRIHAASVWVGWTFPA 496
505 FIVAFIPLSNKEFFKNYGTNGVCEPLHSEDTESIGAOIYSVAIFLGINLAFTIIVSV 564
497 FAALFPL-----FGISSYMKVSIQPM--DIDSPLSQLYMAL-LVNLVAFVIGCY 548
565 GSMFYSVQASITATEINQVKKEMITAKRFFIYFDALCMIPFVVKFSLLOVEIIG 624
549 THYIYLVNPTIVS-----SSDKIKAKRMATLIFTDFLMAPISPFAISASIKVPL- 601
625 TITSVVVFIL--PINSALNPLVLTITRPKE-----MIHP-----W 661
602 TVSRAKILLVLFYINSCANPFLYAITKRNDRDFILLSKFGCYEMOAIYRTTSSAT 661
662 YNYRQKSMDSKQKTYAPFETIWEV 687
662 HNFARKSHGSSAPRV-TNSYLVPL 686

RESULT 9
LSHR_PIG STANDARD; PRT; 696 AA.
AC P16582;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lutropin-choriogonadotropic hormone receptor precursor (LH/CG-R)
DE (LSH-R) (lutinizing hormone receptor).
GN LHGR.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=69332517; PubMed=2502844;
RA Looftelt H., Mistran M., Alger M., Salese R., Thi M.T.V.H.-L.,
RA Jolivet A., Guichon-Mantel A., Sar S., Jallat B., Garnier J.,
RA Milgrom E.;
RT "Cloning and sequencing of porcine LH-hCG receptor cDNA: variants
RT lacking transmembrane domain.";
RL Science 245:525-528(1989)
CC -1- FUNCTION: RECEPTOR FOR LUTROPIN-CHORIOGONADOTROPIC HORMONE.
CC THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH
CC ACTIVATE ADENYLATE CYCLASE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- ALTERNATIVE PRODUCTS: 4 ISOFORMS: A (SHOWN HERE), B, C AND D; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC FSH/USH/TSH SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M29525; AAA31062.1; -;
DR EMBL; M29526; AAA31063.1; -;
DR EMBL; M29527; AAA31064.1; -;
DR EMBL; M29528; AAA31065.1; -;
DR PIR; A41344; A41344.
DR PIR; B41344; B41344.
DR PIR; C41344; C41344.
DR PIR; D41344; D41344.
DR HSSP; P22888; ILUT.
DR GCRDB; GCR_0106; -;
DR GCRDB; GCR_0608; -;
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR00372; LRR_Nterm.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR; 2.
DR PRINTS; PRO0373; GLYCOPROTEIN.
DR PRINTS; PRO1144; LSHRECEPTOR.
DR SMART; SM00013; LRRNT; 1.
DR PROSITE; PS00237; G-PROTEIN_RECPEP_FL_1; 1.
DR PROSITE; PS0262; G-PROTEIN_RECPEP_FL_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Phosphorylation; Repeat; Leucine-rich repeat; Alternative splicing.
FT SIGNAL 1 27
FT CHAIN 28 696
FT FT
FT DOMAIN 28 358
FT TRANSSEM 359 386
FT DOMAIN 387 395
FT TRANSSEM 396 418
FT DOMAIN 419 439
FT TRANSSEM 440 462
FT DOMAIN 463 482
FT TRANSSEM 483 505
FT DOMAIN 506 525
FT TRANSSEM 526 547
FT DOMAIN 548 570
FT TRANSSEM 571 594
FT DOMAIN 595 605
FT TRANSSEM 606 626
FT DOMAIN 627 696
FT REPEAT 48 71
FT REPEAT 122 147
FT REPEAT 149 171
FT LRR 1.
FT LRR 2.
FT LRR 3.

```

FT REPEAT 172 196 LRR 4.
FT REPEAT 198 220 LRR 5.
FT REPEAT 221 244 LRR 6.
FT DISULFID 439 514 BY SIMILARITY.
FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 195 195 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 291 291 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 313 313 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VASAPLIC 317 329 YSAIFASELSDW -> LILGALPATRCLS (IN ISOFORM B).
FT VASAPLIC 330 696 MISSING (IN ISOFORM B).
FT VASAPLIC 317 331 YSAIFASELSDMY -> SKSRADGYQQRHKDC (IN ISOFORM C).
FT VASAPLIC 332 696 MISSING (IN ISOFORM C).
FT VASAPLIC 317 628 MISSING (IN ISOFORM D).
SQ SEQUENCE 696 AA; 78092 MW; 593DEF1C25F982FE CRC64;

```

Query Match 12.68; Score 481.5; DB 1; Length 696;
 Best Local Similarity 23.38; Pred. No. 3.7e-22;
 Matches 144; Conservative 133; Mismatches 261; Indels 79; Gaps 17;

```

OY PSVSNVTAMSLQNNLIRKLPDCKKNYHDLQKLDL-ONNKITSISIAVFGINSITKLY 145
DB 45 PGPRAGLSRLLTLPKIVIPSOAFRGINEVYKLEISQSDLEKTEANAFDNLNLSEL 104
OY 146 LSHNR-ITFLPGVEEDLRLLEWLEIEDNHLKRISPT-FYGLNSLILVLMNVLTPL 202
DB 105 IONKKNLYIIBPGAFATNLPRLKYLISCTGTRKLPDVKRISSENFLEICDNHITTV 164
OY 203 PDKPLCQMPPLHMLDLEGNIHNLNLTFTSCSNLTVLVARKKINLNLNENTAPLOKL 262
DB 165 PANAFOGANNESITLKLTVNGEELIOSHAFNGTLLISLEKENALKKMHNDAPFGARGP 224
OY 263 DELDGSKNKINLPLPIFKDEL---SQLNLSYPIQKIQANQDQVYKLSLEGLIE 319
DB 225 SILDSSSTKIDALPSYGLSISOTLATISYSKLKLPREKFTNLDAVLTPSHCC----- 280
OY 320 ISNIQORMRPL---MNLSHYFKKQ----- 343
DB 281 -----AFRNLPTKEQNFSEIFKFNKQCESTARPNNETLSAIFASELSDMDYD 333
OY 344 -YCGAAPHVRCCKPTDQSSISLENLAIIQRYVFWVVSATCGENITVYICRPIYRSEN 402
DB 334 GFC--SPKTLQCAPRPPDAFNCEDIMGYDFLRVLIMLILNLAIMGWVTLFV--LLTSHY 389
OY 403 KLYAMSI--SLCCADCLMGITLFEVIGDFLFRGEYKKNHQAOLMWESTHCOLVSLATLS 460
DB 390 KLYPRFLMCLNLSFADPCMGILLLILIASVDAQTKQYNNHAIIDMOTGCGSVAAGFTYFA 449
OY 461 TEVSALLITFLLEKYICIVYPRFCVRPKCR-TITVLILWITGFIYAFIPLSKKEFK 519
DB 450 SELSYVTLTVTLERWHTITTAIQDOKRLRHAIRPIMLGWLFSTLAMLPLVG-----V 505
OY 520 NYGNGVCFPLHSEDTSSIAQIYVAIFLCINAAFIIVFSGSFYSYHOSAITAT 579
DB 506 SSYKVSICLPM--DVEFTTISOYVITLIL-LNVVAFIILICACIKIFYFAVQNDLMAT 561
OY 580 EIRNOVKEMILAKREFEIVFTDALCWPVIVVVKFLSLQVEIPIGTITSVVITL--PI 637
DB 562 -----NKDTIAKKMAVLITFDTCMAPISFALSALKVPLI-TVINSKVLVLVLEFPV 614
OY 638 NSALNPILYITLTTRPK 654
DB 615 NSCANPFLYIAFTKAFR 631

```

RESULT 10
 FSHR_MACFA STANDARD; PRT; 695 AA.
 ID FSHR_MACFA
 AC P32212;

```

DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Follicle stimulating hormone receptor precursor (FSH-R) (Folliotropin receptor).
GN FSHR.
OS Macaca fascicularis (Crah eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Macaca.
OC Cercopithecoidea; Macaca.
OC NCBI_Taxid=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=94071854; PubMed=7504463;
RA Gromoll J., Dankbar B., Sharma R.S., Nieschlag E.;
RT "Molecular cloning of the testicular follicle stimulating hormone receptor of the non human primate Macaca fascicularis and identification of multiple transcripts in the testis.";
RL Biochem. Biophys. Res. Commun. 196;1066-1072(1993).
CC -!- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLYLATE CYCLASE.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS. FSH/LSH/TSR SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC
CC EMBL; X74454; CAA52463.1; -.
DR PIR; S36452; S36452.
DR PIR; JN0898; JN0898.
DR HSSP; P23945; IXUN.
DR GCRDB; GCR_0653; -.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000372; LRR_Nterm.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR; 4.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PR00373; GLYCROMONER.
DR PRINTS; PR01143; FSHRECEPTOR.
DR SMART; SM00013; LRRNT; 1.
DR PROSITE; PS00237; G-PROTEIN_RECEP_FL_1; 1.
DR PROSITE; PS0262; G-PROTEIN_RECEP_FL_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal; Phosphorylation; Repeat; Leucine-rich repeat.
FT SIGNAL 1 17
FT CHAIN 18 695
FT DOMAIN 18 366
FT TRANSSEM 367 387
FT DOMAIN 388 398
FT TRANSSEM 399 421
FT DOMAIN 422 443
FT TRANSSEM 444 465
FT DOMAIN 466 485
FT TRANSSEM 486 508
FT DOMAIN 509 528
FT TRANSSEM 529 550
FT DOMAIN 551 573
FT TRANSSEM 574 597
FT DOMAIN 598 608
FT TRANSSEM 609 630
FT DOMAIN 631 695
FT REPEAT 44 68
FT REPEAT 69 93
FT LRR 1.
FT LRR 2.

```


RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bona M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzatelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasakawa H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Welter C., Whitaker C., Wilming L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.,
 RA "Functional annotation of a full-length mouse cDNA collection.",
 RL Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE OF 1-51 FROM N.A.
 RX MEDLINE=93093308; PubMed=1459341;
 RA Hultainen I.T., Eskola V., Pakarinen P., Matikainen T.,
 RA Sprengel R.;
 RT "The murine luteinizing hormone and follicle-stimulating hormone
 RT receptor genes: transcription initiation sites, putative promoter
 RT sequences and promoter activity.";
 RL Mol. Cell. Endocrinol. 88:55-66(1992).
 CC -1- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY
 CC OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
 CC ADENYLATE CYCLASE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC FSH/LSH/TSH SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AF095642; AAC67559.1; -;
 DR EMBL: AK016635; BAB30351.1; -;
 DR EMBL: S49632; AAB24401.1; -;
 DR EMBL: M87570; AAA37641.1; -;
 DR GCRDP: GCR.0304; -;
 DR MGD: MGI:95583; Fshr.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR003372; LRR_Nterm.
 DR Pfam: PF00001; 7tm.1; 1.
 DR Pfam: PF00560; LRR; 3.
 DR Pfam: PF01462; LRRNT; 1.
 DR PRINTS: PR00237; GPCR_RHODOPSIN.
 DR SMART: SM00013; LRRNT; 1.
 DR PROSITE: PS00237; G-PROTEIN_RECP_F1_1; 1.
 DR PROSITE: PS00262; G-PROTEIN_RECP_F1_2; 1.
 KM G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KM Phosphorylation; Repeat; Leucine-rich repeat.
 FT SIGNAL 1 17
 FT CHAIN 18 692
 FT DOMAIN 18 365
 FT TRANSMEM 366 386
 FT DOMAIN 387 397
 FT TRANSMEM 398 420
 FT DOMAIN 421 442
 FT TRANSMEM 443 464
 FT DOMAIN 465 507
 FT TRANSMEM 508 527
 FT TRANSMEM 528 549
 FT TRANSMEM 550 572
 FT TRANSMEM 573 596
 FT DOMAIN 597 607
 FT TRANSMEM 608 629
 FT DOMAIN 630 692
 FT REPEAT 44 68
 FT REPEAT 69 93

FT REPEAT 119 143 LRR 3.
 FT REPEAT 170 192 LRR 4.
 FT REPEAT 193 216 LRR 5.
 FT REPEAT 218 240 LRR 6.
 FT DISULFID 441 516 BY SIMILARITY.
 FT CARBOHYD 191 191 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 436 436 O -> K (IN REF. 2).
 SQ SEQUENCE 692 AA: 77769 MW: 485722918056344 CRC64;
 Query Match 12.4%; Score 475.5; DB 1; Length 692;
 Best Local Similarity 23.5%; Pred. No. 8,5e-22;
 Matches 162; Conservative 125; Mismatches 265; Indels 137; Gaps 25;
 QY 35 GNTTKLPQLHCGVDDCGQADEDCVYVLCQMSLPGELDMMKRFVSVSSTVT 94
 DB 13 GSGGCHHMLCHCSN-----RVFLCDSKVTEIP-----PDLPRNAI 49
 QY 95 AMSLQWNLIRKLPPDCFRNYHDLOKLDL-ONNKITISIVAFRGNSLTKYLSH-NRIT 152
 DB 50 ELRFVLTAKRYTPKGSFSGFDELEKIEISQNDYLEVADYFSLPNIHEIRKANNL 109
 QY 153 FLKPGVEDLRLMLIEDNHLRISPPPEYGLNSL-ILLVLMNVYLRPLDPKPLCOH 210
 DB 110 YINFEARQNPISLRYLISNTGIKHL-PAFHKIQSLQKVLDDIQDN----- 155
 QY 211 MPRLMHLDLSGNHINHLNLTFLSCS-NLYLVYVRKKKINHLNENTRAPLOKDELDG- 268
 DB 156 -----NHIILARNSFMGLSFESVYLMANKNGIOGIRHCAANGTO-IDELNLSID 202
 QY 269 SNKIENPLPIFKD-----LEKLSQNLNSYNDPIQIQAQNDYVYKLSLS-L 315
 DB 203 NNNEELPDDYFGASGPVYLDISRTKYSLPNIENLEKLRARSRYRLKPLSLDKFV 262
 QY 316 EGIEIS-----NIQRMFRPLMNL----- 334
 DB 263 MLIRASLTPYPSHCAPANMWRQTSHELPHCKSISRODIDDMTPGDQRYSLVDDEPSYG 322
 QY 335 --SHIYFKKQY--CGAAPHYRCKPMTDGISSLENLASSIQVFWVYVSAVCFENIF 390
 DB 323 KGSQMLYSEDDYDLCNFEVDY-TCSPKPDAFNPCEIDMGVILVILWFISLAIITNTT 381
 QY 391 VICRPIYRSENKLYAMSI--SLCCADCLMGILYFYIGFDELFKREYKNHAOLMEST 448
 DB 382 VLWY--LTSQYKLTFRPLMNCNLAFLADLCIGYLLILLASVDHTKQYHVAIDMPTGA 439
 QY 449 HQQLVGSIALISTEVSVLLTFLTELYKICIVYPRC-VRPGRCTTVLLIMVIGFIV 507
 DB 440 GCDAAGFYVFASELVYTLAATILERMHTTHAMQLECKVQLCHASIMVLCWAFAPAA 499
 QY 508 AFPLSKKEPKKNYNGVNGCFPLHSDTESIGAOISVAMIFGLINLAFFIIFSGSM 567
 DB 500 ALFPI-----FGISYMKVYSICLPW---DIDPSLSQLYVMA-LVLNLAFAVIGCTTHI 551
 QY 568 FYSYHOSAITATEIRNOYKKEMLAKRFFVFETDACPWFVFKFLSLJOYEIPGTIT 627
 DB 552 YLYVRNNIVSS-----SBDTKAKRMATLITFDLCMPILFEAFASLSKVLPI-TVS 604
 QY 628 SWVYVIL--PINSALNPILYTLTPPK 654
 DB 605 KAKILIVLFYPIVNSCANPFLAIFTKMR 633
 RESULT 13
 FSHR_EQDAS STANDARD; PRT; 687 AA.
 ID FSHR_EQDAS
 AC 095179;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Follicle stimulating hormone receptor precursor (FSH-R) (Follictropin

DE receptor).
 GN FSHR.
 OS Equus asinus (Donkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 OX NCBI_TaxID=9793;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Testis;
 RX MEDLINE=97338913; PubMed=9195473;
 RA Richard F., Martinat N., Remy J.-J., Salese R., Combarnous Y.;
 RT "Cloning, sequencing and in vitro functional expression of
 RT recombinant donkey follicle-stimulating hormone receptor: a new
 RT insight into the binding specificity of gonadotrophin receptors.";
 RL J. Mol. Endocrinol. 18:193-202(1997)
 CC -1- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY
 CC OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
 CC ADENYLATE CYCLASE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC FSH/LSH/TSH SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U73659; AAB18245.1; -.
 DR HSSP: P23945; 1XUN.
 DR GCRdb: GCR_1116; -.
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR00372; LRR_Nterm.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 4.
 DR Pfam: PF01462; LRRNT; 1.
 DR PRINTS: PRO0237; GPCRHRHODPSN.
 DR SMART: SM00013; LRRNT; 1.
 DR PROSITE: PS00237; G-PROTEIN_RECPT_FL1; 1.
 DR PROSITE: PS0262; G-PROTEIN_RECPT_FL2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Phosphorylation; Repeat; Leucine-rich repeat.
 FT SIGNAL 1 17
 FT CHAIN 18 687 FOLLICLE STIMULATING HORMONE RECEPTOR.
 FT DOMAIN 18 358 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 359 379 1 (POTENTIAL).
 FT DOMAIN 380 390 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 391 413 2 (POTENTIAL).
 FT DOMAIN 414 435 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 436 457 3 (POTENTIAL).
 FT DOMAIN 458 477 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 478 500 4 (POTENTIAL).
 FT DOMAIN 501 520 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 521 542 5 (POTENTIAL).
 FT DOMAIN 543 565 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 566 589 6 (POTENTIAL).
 FT DOMAIN 590 600 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 601 622 7 (POTENTIAL).
 FT DOMAIN 623 687 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 44 68 LRR 1.
 FT REPEAT 69 93 LRR 2.
 FT REPEAT 119 143 LRR 3.
 FT REPEAT 170 192 LRR 4.
 FT REPEAT 193 216 LRR 5.
 FT REPEAT 218 240 LRR 6.
 FT DISULFID 434 509 BY SIMILARITY.
 FT CARBOHYD 191 191 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT

SQ SEQUENCE 687 AA; 76937 MW; FC3AF0B5531DA9A CRC64;
 Query Match 12.4%; Score 474.5; DB 1; Length 687;
 Best Local Similarity 24.1%; Pred. No. 9,7/e-22;
 Matches 161; Conservative 126; Mismatches 255; Indels 127; Gaps 24;
 QY 53 CGNQADENCVYVLCQCMSPGLGELDMKPKPTSVS-VSSNVTAMSLDMNLRKPPPCF 111
 DB 18 CHQVCHYSNRFVLCQ-----ESKVEYEDSDLPRLALEFVLYKLVYKCAAF 66
 QY 112 KNYHDLQKLDLNNKNTSISYAFKGLNSLTKLYLSHNRITFLKPGVEDLRLLEWLTIE 171
 DB 67 SGFGDLKKEIEISQNDVLEV-----IEANVSNPKLKEIRIE 103
 QY 172 D-NHLSRISPPFEGVGLNSLILVLMNNVLYTRLPDKPLQCMHRLWDLQGN-HIHNIRN 229
 DB 104 KANMLLYIDHDAFQNLPMLOYLLISNTGICKLPVAKKIQSLQKV-LDIDQNNINIHVER 162
 QY 230 LRFISCS-NLNVLYVRKKKINHLNENTFAPLOKIDELDG-SNKIENLPPIFKDLKELS 287
 DB 163 NSFGLSESMILRLSKNGIQEIHNCAPNGTQ-LDELWLSDDNNLEELPNDVFGASGPV 221
 QY 288 QUNLSYNEIQRKIQANQFVLYVLYKLSLEGI-EISNIOQRMRPPLMNLSHY----- 338
 DB 222 ILDISGTRHSLPNVGLNKLKRLARSTYNLKLPSLEK--FVALMEASLTPSHCAFA 279
 QY 339 -----FKKQY--CGYAPHVASC 354
 DB 280 NWROQTSELQTCNKSILROEVDMTQARGERVSLAEDDESMMYSSEFDYDLCNEVDV-TC 338
 QY 355 KPNNDGISLENLASIIQRFVFWVAVTQCGNIFVCMRPYISENKLAMSLI--SL 412
 DB 339 SPKPDAPNCPEDIMGYDLRLVLMFISLITATGNIIVYT--LITSQYKLVPRRLKCNL 396
 QY 413 CCADCLMGIVYLVIGGFDLKEGEXYKKAQOLMESTHCOLVSLAISTEVSALLTFELT 472
 DB 397 AFADLCIGIYLLIASVYDIHRSQYHNVAIDMQTAGDAGGFVFPSELSTVYLTAIT 456
 QY 473 LEKYCIYPPRCVRRPKCR-----TIVLLIIVITGTVAFIPLSNKEFFKNYGTNGV 527
 DB 457 LERWHTITHAMQL-----ECKYQLRHAASVMYGVFGVGVGLPL-----FGISTYMKYSI 508
 QY 528 CEPHLSEPTESGAIQYSAVAFGLINLAFTIIVSYSGSMFYSVQSAITAEINQYK 587
 DB 509 CLPM--DIDPSLSQLYMSL-LVLNLAFAVYICGYTHIYLVNPNIVS-----SS 558
 QY 588 EMILAKRFFEFYFDALCMIPFVVKFSLDVEIPGTTSMVVFIL--PINSALNPL 645
 DB 559 DTKIKRKGILITFDPLCMAPISFSGISASLKVALLI-IVSKSKILLVLFYPINSCANFPL 617
 QY 646 YTLTRPPK 654
 DB 618 YAIPTKNER 626
 RESULT 14
 LSHR_MOUSE STANDARD; PRT; 700 AA.
 ID LSHR_MOUSE
 AC P30730;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lutropin-choriogonadotropin hormone receptor precursor (LH/CG-R)
 DE (LSH-R) (Luteinizing hormone receptor).
 DE LHCGR OR LHR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92165799; PubMed=1311310;

FT	CARBOHYD	103	103	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	178	178	N-LINKED (GLCNAC. . .) <th>(POTENTIAL).</th>	(POTENTIAL).
FT	CARBOHYD	199	199	N-LINKED (GLCNAC. . .) <th>(POTENTIAL).</th>	(POTENTIAL).
FT	CARBOHYD	295	295	N-LINKED (GLCNAC. . .) <th>(POTENTIAL).</th>	(POTENTIAL).
FT	CARBOHYD	303	303	N-LINKED (GLCNAC. . .) <th>(POTENTIAL).</th>	(POTENTIAL).
FT	CARBOHYD	317	317	N-LINKED (GLCNAC. . .) <th>(POTENTIAL).</th>	(POTENTIAL).
SO	SEQUENCE	700 AA;	78214 MM;	8A6840A011E1E014	CRC64;

Query Match	12.3%	Score 471.5;	DB 1;	Length 700;
Best Local Similarity	23.8%	Pred. No. 1.5e-21;		
Matches 148;	Conservative 131;	Mismatches 271;	Indels 71;	Gaps 18;

QY	87	PSVSNVTAMSLQWNLILKRLPPDCFKNHDQKIDL-ONNKITSISIAFAGLSNSTKLY	145
QY	49 <td>PGRPRAGLARSLLTFLPVAVIPSQAFRGINTEVVKIEISQSDSLERLEANAFFNLNLSSIL<td>108</td></td>	PGRPRAGLARSLLTFLPVAVIPSQAFRGINTEVVKIEISQSDSLERLEANAFFNLNLSSIL <td>108</td>	108
QY	146	LSHNR-ITFLKPGVPEDLHLEWLIIEDN-----HLSRISPPFFYGLNSLLVLMNV <td>198</td>	198
DB	109	IQNKNLIIIEPGAFTMLPRKLYSLICGTGTRLPDSKISSSEF---NFLEICDLY <td>164</td>	164
QY	199	LTRLPDKPLCOHMRPLHMLDLEGNIHMLRLMTFTSCSNLVLVWRKKIN---- <td>HLNEN 254</td>	HLNEN 254
DB	165	ITTPGNAFGQMNNESTLKLXGNFEVQSHAF-----NGTLLSLBELKENIYLEKMHSG	220
QY	255	TFAPRLKIDELDLGSKNTENP-----PLIFKDLKLSQL-----NLST-	293
DB	221	TFQCATGPIILDVSTKQALPSHGLEISQTLIATSSYSLKTLPSREKFTSLVATLYP	280
QY	294	-----NPQIKQANOPDYLVLKLSLSLEGIEISNIOQRFRPLMNSHYPFKFOYCG	346
DB	281	SHCCAFRLPKRQGNFSTISFENSKEQESYVREAN-NEITXSAFEEENELSGMDYIDF	339
QY	347	YAPHVRSKPRTDGISLLENLAIIOREFWVAVYSAVTCFNGINFIYICMPYIRSEKLYA	406
DB	340	CSPKTLQCTPEPDANPCEDIMGYAFVLVLIIMLIIAIFGNLVFLVFLTSRYKLVPR	399
QY	407	MSIISLCCADCLMGITLVFVIGGFDLKFGEYENKHAQLMESTHOQVSLAISLEVSYL	466
DB	400	FLMCLNSFADCMGILYLLIASVDSQTKGYUHNADIMQTSGSGAMGEFTVFVASELSY	459
QY	467	LATFELTEKYLCIYPPFCVARGKCR-TTYVLIIIMINGFIVAEIPLSNKEFFKNYGTN	525
DB	460	TLVYTLIERMHTITTYAVGDLQKLRHRIAPIMLGWIPSTLMTPLVG---VSSYKVV	515
QY	526	GVCPRPLSHEDIESGAQIYSVAIFGLINLAFTIIVESYSGMFSYRQSAITAREINOV	585
DB	516	SICLPM-----DVESTLSQYIILSILL-LNAVAFVVIACVYVAVAYONDELTA-----	565
QY	586	KKEMLIAKRFPIVETDALCWPITFVVVFSLLDVEIRGTTSWVPIRL--PINSALNP	643
DB	566	NKDKRIAKKMAILFTDPTCMAPISFAISAFVPLP-ITYNSKVLVLLVPPYNSCANP	624
QY	644	ILYTLITRPEKE---MIMRF	660
DB	625	FLYAVTFKAFQDFELLSP	645

RESULT 15	LSHR_RAT	STANDARD;	PRT;	700 AA.
AC	P16235; P70646; Q63807; Q63808; Q63809;			
DT	01-Apr-1990 (Rel. 14, Created)			
DT	01-Apr-1990 (Rel. 14, Last sequence update)			
DT	16-Oct-2001 (Rel. 40, Last annotation update)			
DE	Lutropin-Choriogonadotropic hormone receptor precursor (LH/CG-R)			
DE	(LSH-R) (Luteinizing hormone receptor).			
GN	LHCGR.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
IN	[1]			

RP SEQUENCE FROM N.A.
RA MEDLINE=89332512; PubMed=2502842;
RX Mefarland K.C., Sprengel R., Phillips H.S., Koehler M.,
RA Rosembli N., Nikolics K., Segaloff D.L., Seeburg P.H.;
RT "Lutropin-choriogonadotropin receptor: an unusual member of the G
RL protein-coupled receptor family."; Science 245:494-499(1989).
RN [2]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Ovary;
RX MEDLINE=92347604; PubMed=1353463;
RA Aatsinki J.T., Pietila E.M., Lakkakorpi J.T., Rajaniemi H.J.;
RT "Expression of the LH/CG receptor gene in rat ovarian tissue is
regulated by an extensive alternative splicing of the primary
transcript"; Mol. Cell. Endocrinol. 84:127-135(1992).
RL [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=91209270; PubMed=2019252;
RA Koo Y.B., Slaughter R.G., Ji T.H.;
RT "Structure of the luteinizing hormone receptor gene and multiple
exons of the coding sequence."; Endocrinology 128:2297-2308(1991).
RN [4]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=91006819; PubMed=1976554;
RA Bernard M.P., Myers R.V., Moyle W.R.;
RT "Cloning of rat lutropin (LH) receptor analogs lacking the soybean
lectin domain."; Mol. Cell. Endocrinol. 71:R19-R23(1990).
RL [5]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=91126285; PubMed=2281186;
RA Segaloff D.L., Sprengel R., Nikolics K., Ascoli M.;
RT "Structure of the lutropin/choriogonadotropin receptor."; Recent Prog. Horm. Res. 46:261-303(1990).
RL [6]
RP SEQUENCE OF 295-700 FROM N.A.
RX MEDLINE=91060531; PubMed=2174034;
RA Tsai-Morris C.H., Buckzo E., Wang W., Dufau M.L.;
RT "Intronic nature of the rat luteinizing hormone receptor gene defines
a soluble receptor subspecies with hormone binding activity."; J. Biol. Chem. 265:19385-19388(1990).
RL [7]
RP SEQUENCE OF 27-37.
RX MEDLINE=89174723; PubMed=2925659;
RA Roche P.C., Ryan R.J.;
RT "Purification, characterization, and amino-terminal sequence of rat
ovarian receptor for luteinizing hormone/human chorionadotropin."; J. Biol. Chem. 264:4636-4641(1989).
RL [8]
RP MUTAGENESIS.
RX MEDLINE=91332007; PubMed=1714448;
RA Ji I., Ji T.H.;
RT "Asp383 in the second transmembrane domain of the lutropin receptor
is important for high affinity hormone binding and cAMP production."; J. Biol. Chem. 266:14953-14957(1991).
RL [9]
RP FUNCTION: RECEPTOR FOR LUTROPIN-CHORIOGONADOTROPIC HORMONE.
CC THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH
CC ACTIVATE ADENYLYLATE CYCLASE.
CC [10]
CC SUBCELLULAR LOCATION: Integral membrane protein.
CC [11]
CC ALTERNATIVE PRODUCTS: AT LEAST 11 ISOFORMS WHICH DIFFER IN
CC SUBCELLULAR LOCATION ARE PRODUCED BY ALTERNATIVE SPLICING
CC OF THE SAME GENE.
CC [12]
CC SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC [13]
CC SIMILARITY: CONTAINS 7 LEUCINE-RICH REPEATS (LRR).
CC [14]
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M26199; AAA41528.1; -
DR EMBL; M61212; AAA41527.1; -
DR EMBL; M61211; AAA41527.1; JOINED.
DR EMBL; S40803; AAB22680.1; -
DR EMBL; S40787; AAB22680.1; JOINED.
DR EMBL; S40903; AAB22680.1; JOINED.
DR EMBL; S40904; AAB22680.1; JOINED.
DR EMBL; S40905; AAB22680.1; JOINED.
DR EMBL; S40907; AAB22680.1; JOINED.
DR EMBL; S40909; AAB22680.1; JOINED.
DR EMBL; S40918; AAB22680.1; JOINED.
DR EMBL; S40920; AAB22680.1; JOINED.
DR EMBL; S40795; AAB22680.1; JOINED.
DR EMBL; S40798; AAB22680.1; JOINED.
DR EMBL; S40799; AAB22680.1; JOINED.
DR EMBL; S40787; AAB22681.1; JOINED.
DR EMBL; S40903; AAB22681.1; JOINED.
DR EMBL; S40904; AAB22681.1; JOINED.
DR EMBL; S40905; AAB22681.1; JOINED.
DR EMBL; S40907; AAB22681.1; JOINED.
DR EMBL; S40909; AAB22681.1; JOINED.
DR EMBL; S40918; AAB22681.1; JOINED.
DR EMBL; S40920; AAB22681.1; JOINED.
DR EMBL; S40803; AAB22682.2; -
DR EMBL; S40787; AAB22682.2; JOINED.
DR EMBL; S40903; AAB22682.2; JOINED.
DR EMBL; S40904; AAB22682.2; JOINED.
DR EMBL; S40905; AAB22682.2; JOINED.
DR EMBL; S40907; AAB22682.2; JOINED.
DR EMBL; S40909; AAB22682.2; JOINED.
DR EMBL; S40918; AAB22682.2; JOINED.
DR EMBL; S40920; AAB22682.2; JOINED.
DR EMBL; S40795; AAB22682.2; JOINED.
DR EMBL; S40798; AAB22682.2; JOINED.
DR EMBL; S40799; AAB22682.2; JOINED.
DR EMBL; S40787; AAB22683.1; JOINED.
DR EMBL; S40903; AAB22683.1; JOINED.
DR EMBL; S40904; AAB22683.1; JOINED.
DR EMBL; S40905; AAB22683.1; JOINED.
DR EMBL; S40907; AAB22683.1; JOINED.
DR EMBL; S40909; AAB22683.1; JOINED.
DR EMBL; S40918; AAB22683.1; JOINED.
DR EMBL; S40920; AAB22683.1; JOINED.
DR EMBL; S40795; AAB22683.1; JOINED.
DR EMBL; S40798; AAB22683.1; JOINED.
DR EMBL; S40799; AAB22683.1; JOINED.
DR EMBL; S40803; AAB22684.2; -
DR EMBL; S40787; AAB22684.2; JOINED.
DR EMBL; S40903; AAB22684.2; JOINED.
DR EMBL; S40904; AAB22684.2; JOINED.
DR EMBL; S40905; AAB22684.2; JOINED.
DR EMBL; S40907; AAB22684.2; JOINED.
DR EMBL; S40909; AAB22684.2; JOINED.
DR EMBL; S40918; AAB22684.2; JOINED.
DR EMBL; S40920; AAB22684.2; JOINED.
DR EMBL; S40795; AAB22684.2; JOINED.
DR EMBL; S40798; AAB22684.2; JOINED.
DR EMBL; S40799; AAB22684.2; JOINED.
DR EMBL; M68928; AAA41529.1; -
DR EMBL; M68917; AAA41529.1; JOINED.
DR EMBL; M68918; AAA41529.1; JOINED.
DR EMBL; M68919; AAA41529.1; JOINED.
DR EMBL; M68920; AAA41529.1; JOINED.
DR EMBL; M68921; AAA41529.1; JOINED.
DR EMBL; M68922; AAA41529.1; JOINED.
DR EMBL; M68923; AAA41529.1; JOINED.
DR EMBL; M68925; AAA41529.1; JOINED.
DR EMBL; M68926; AAA41529.1; JOINED.
DR EMBL; M68927; AAA41529.1; JOINED.
DR PIR; A32460; A32460.
DR PIR; A41343; A41343.
DR HSP; P22886; 1IUT.
DR GCRDb; GCR_0138; -
DR GCRDb; GCR_0139; -
DR GCRDb; GCR_0262; -


```

DR GCRD; GCR_0612; -
DR GCRD; GCR_0613; -
DR GCRD; GCR_0614; -
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR InterPro: IPR001611; LRR.
DR Pfam: PF00001; 7tm_1; 1.
DR Pfam: PF00560; LRR; 2.
DR PRINTS: PRO0373; GLYCHORMONER.
DR PRINTS: PRO1144; LSHRECEPTOR.
DR SMART: SM00013; LRNT; 1.
DR PROSITE: PS00237; G-PROTEIN_RECP_F1_1; 1.
DR PROSITE: PS00262; G-PROTEIN_RECP_F2_1; 1.
DR G-protein coupled receptor; Transmembrane; signal;
KW Phosphorylation; Repeat; Leucine-rich repeat; Alternative splicing.
FT STGNL 1 26
FT CHAIN 1 700
FT DOMAIN 27 362
FT TRANSMEM 363 390
FT DOMAIN 391 399
FT TRANSMEM 400 422
FT DOMAIN 423 443
FT TRANSMEM 444 466
FT DOMAIN 467 486
FT TRANSMEM 487 509
FT DOMAIN 510 529
FT TRANSMEM 530 551
FT DOMAIN 552 574
FT TRANSMEM 575 598
FT DOMAIN 599 609
FT TRANSMEM 610 631
FT DOMAIN 632 700
FT REPEAT 52 75
FT REPEAT 126 150
FT REPEAT 152 175
FT REPEAT 176 200
FT REPEAT 202 224
FT REPEAT 225 248
FT REPEAT 250 271
FT DISULFID 443 518
BY SIMILARITY.

```

Query Match 12.3%; Score 470.5; DB 1; Length 700;
 Best local similarity 22.8%; Pred. No. 1,7e-21;
 Matches 149; Conservative 135; Mismatches 233; Indels 137; Gaps 23;

```

QY 87 PVSANVTAMSLQNLIRKLPDPCFKNYHDLQKIDLQNNKITSISYARGLNSLTSLKLY 145
DB 49 GPPRAGLARSLTYLPVKAIPSAQAFGLNEVYKIEISQDSLRIEANAFAFDNLNLSEL 108
QY 146 LSHNR-ITFLKPGVEEDLHLREMLIEDN-----HLSRISPTFFYGLNSLILVLMNV 198
DB 109 IONTKNLITIEPGAFTLPLRKLYSLICNTGRTLPDYTKISSSEF-----NFILEICNLH 164
QY 199 LTRLDPKPLCQHMRLHMLDEGNHINLNLNLFISCSNLTVLVMRKKNINHLNENTFAP 258
DB 165 ITTIPGNAFOGNNESVTLYKYGNGFEVQSHAF-----NGTTLISLE-----LKENIY-- 213
QY 259 LQKIDE-----LDGSKNIENLPLRIKDLKELSQLNSYNPIQKIQANOPDY 307
DB 214 LEKHSAGAFGANGPSILDISSTKLALPS-----HGLESIQF-----L 252
QY 308 VKLSLSLEGLEISNIQRMERPLM-----NLSHYFKKQ--- 343
DB 253 IALSSYSKTLTP-----SKKFTSLVATLTYPSHCACFRNLPKKEONFSFSEFENFSKQC 308
QY 344 -----YCGYAPHVRSCKPNTDGISSLENLALSIQ 373
DB 309 ESTYRKADNETIYSAIPEENNELSGWDYDGC-SPRTIQCAPEDAFNCPCEDIMGYAFL 366
QY 374 RVEYVWVSAVTCGNIIVICMRPIYRSEKLYAMSIISLCCADCLMGIVLFTVIGFDLKF 433
DB 367 RVLIMILNIIAIFGNLVLEFLVLTLSRYKLTVPRLMCNLSFADFQCMGLAYLLILASVDSQT 426

```

```

QY 434 RGEYKHAQIMMESTHCOLVSLAISTEVSULLFTLEKYICIVYPPRCVPRGKCR- 492
DB 427 KGVYNNAIIDMVGSGGGAAGFTVFSELSVYTLITYTLERHITTYAVQDLQKRLRH 486
QY 493 TITVLLIWTGFIYAFIPISNKEEFKNYGTNGVCPPLHSEDTESIGAOIYSAIPLGI 552
DB 487 AIPMLGMLFSTLIATPLVG---ISNMKVS-ICLPM--DVESTLSQVYLISILI-L 538
QY 553 NIAFIITVSYSGMEFSVHOSATATATEINQVKKEMILAKRFFIYETDALCMPIFVY 612
DB 539 NVAFVVICAYIRIYFVAVONPELTAP-----NKDKIKAKKAILFTDFTCAPISSEF 592
QY 613 KFLSLQVEIPIGTTISWVIFIL--PINSALNPILYTLTRPERE---MIHRE 660
DB 593 AISAAPKVPPL-IVTNSKILLVLEFPVNSCANPFLVAFITKAFQRFDLILLSR 645

```

Search completed: September 5, 2002, 09:04:06
 Job time: 2325 sec

•
•
•
•

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2002, 08:24:36 : Search time 85.41 Seconds

(without alignments)
1462.385 Million cell updates/sec

Title: US-09-647-067-8

Sequence: 1 MTSGSVFFILIFGKXFSHC.....FTYPCMSLISGSTRLSYS 722

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

- 1: SPREMBL_19:*
- 2: sp.archaea:*
- 3: sp.bacteria:*
- 4: sp.fungi:*
- 5: sp.human:*
- 6: sp.invertebrate:*
- 7: sp.mhcc:*
- 8: sp.organelle:*
- 9: sp.phage:*
- 10: sp.plant:*
- 11: sp.potent:*
- 12: sp.virus:*
- 13: sp.vertebrate:*
- 14: sp.unclassified:*
- 15: sp.rvivirus:*
- 16: sp.bacteriap:*
- 17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3687.5	96.2	757	4 Q9HBX9	Q9hbxb9 homo sapien
2	1891.5	49.3	737	11 Q91Z25	Q91z25 mus musculu
3	599.5	15.6	359	5 Q9VBP0	Q9vbp0 drosophila
4	576	15.0	1012	5 Q95YI6	Q95yi6 asterina pe
5	576	15.0	1280	5 Q95YI7	Q95yi7 asterina pe
6	496	12.9	1050	5 Q9BN18	Q9bn18 drosophila
7	494	12.9	701	13 Q9DGC6	Q9dgc6 oreochromis
8	492.5	12.8	1300	5 Q9NMD6	Q9nmd6 drosophila
9	490	12.8	907	11 Q9Z1P4	Q9z1p4 mus musculu
10	486.5	12.7	1360	5 Q9ND11	Q9nd11 drosophila
11	483.5	12.6	658	13 Q9PVN9	Q9pvn9 oncorhynch
12	481	12.5	693	13 Q9DGC5	Q9dgc5 oreochromis
13	479.5	12.5	696	13 Q9DGF5	Q9dgt5 cynops pyrr
14	478	12.5	662	13 Q9PW16	Q9pw16 clarias gar
15	477.5	12.5	334	5 Q9VYG0	Q9vyg0 drosophila
16	476.5	12.4	907	4 Q75473	Q75473 homo sapien

17	476.5	12.4	951	11 Q9Z2H4	Q9z2h4 rattus norv
18	475	12.4	699	4 Q15996	Q15996 homo sapien
19	472.5	12.3	688	11 Q64183	Q64183 rattus sp.
20	470.5	12.3	907	4 Q9UP75	Q9up75 homo sapien
21	470	12.3	662	13 Q98R84	Q98r84 ictalurus p
22	469	12.2	701	4 Q14751	Q14751 homo sapien
23	467	12.2	951	4 Q9BXD1	Q9bxd1 homo sapien
24	464	12.1	951	4 Q9NYD1	Q9nyd1 homo sapien
25	453	11.8	764	6 Q9BG56	Q9bg56 sus scrofa
26	443	11.6	696	13 Q98T85	Q98t85 ictalurus p
27	443	11.6	793	13 Q91948	Q91948 oncorhynch
28	441.5	11.5	764	11 Q9D697	Q9d697 mus musculu
29	439.5	11.5	763	6 Q9BGH4	Q9bgh4 felis silve
30	439	11.5	739	6 Q9BG55	Q9bg55 sus scrofa
31	428	11.2	814	13 Q91949	Q91949 oncorhynch
32	427.5	11.2	778	13 Q98TF4	Q98tf4 oreochromis
33	425	11.1	779	13 Q91BN7	Q91bn7 morone saxa
34	416.5	10.9	928	4 Q9BYD7	Q9byd7 homo sapien
35	403	10.5	601	13 Q42500	Q42500 melleagris g
36	402	10.5	829	5 Q9VEG4	Q9veg4 drosophila
37	402	10.5	831	5 Q94979	Q94979 drosophila
38	400.5	10.4	673	13 Q90WP8	Q90wp8 podarcis si
39	394	10.3	828	4 Q9HBX8	Q9hbxb8 homo sapien
40	389	10.1	533	4 Q9PIV4	Q9piy4 homo sapien
41	387.5	10.1	724	13 Q9VYP0	Q9vyp0 oncorhynch
42	385	10.0	470	4 Q43200	Q43200 homo sapien
43	373.5	9.7	929	5 Q18759	Q18759 caenorhabdi
44	353	9.2	293	6 Q95MF7	Q95mf7 canis famli
45	340.5	8.9	410	4 Q16225	Q16225 homo sapien

ALIGNMENTS

RESULT	ID	Q9HBX9	PRELIMINARY;	PRT;	757 AA.
Q9HBX9	AC	Q9HBX9:			
DT	01-MAR-2001	(TREMBLrel. 16, Created)			
DT	01-MAR-2001	(TREMBLrel. 16, Last sequence update)			
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)			
DE	LEUCINE-RICH REPEAT CONTAINING G PROTEIN-COUPLED RECEPTOR 7.				
GN	IGR7.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=20388592; PubMed=10935549;				
RA	Hsu S.Y., Kudo M., Chen T., Nakabayashi K., Bhalla A.,				
RA	Van der Spek P.J., Van Duin M., Hsueh A.J.;				
RT	"The three subfamilies of leucine-rich repeat-containing G protein-				
RT	coupled receptors (LGR): identification of LGR6 and LGR7 and the				
RT	signaling mechanism for LGR7."				
RL	Mol. Endocrinol. 14:1257-1271(2000).				
DR	EMBL; ARI90500; AAG17167.1; -				
DR	HSSP; P01130; ILDR.				
DR	InterPro; IPR000276; GPCR_Rhodpsn.				
DR	InterPro; IPR001611; LDL_recept_A.				
DR	InterPro; IPR001611; LRR.				
DR	InterPro; IPR000372; LRR_Nterm.				
DR	InterPro; IPR003592; LRR_out.				
DR	InterPro; IPR003591; LRR_Typ.				
DR	Pfam; PF00001; 7tm_1; 1.				
DR	Pfam; PF00057; Idl_recept_a; 1.				
DR	Pfam; PF00560; LRR_8.				
DR	PRINTS; PR00237; GPCR_RHODOPS.				
DR	SMART; SM00192; LDla; 1.				
DR	SMART; SM00370; LRR_7.				
DR	SMART; SM00013; LRRNT; 1.				
DR	SMART; SM00369; LRR_Typ; 10.				
DR	PROSITE; PS0262; G_PROTEIN_RECPT_FL_2; 1.				

DR PROSITE: PS01209; LDLRA_1; 1.
 DR PROSITE: PS50068; LDLRA_2; 1.
 KW Glycoprotein; Receptor.
 SQ SEQUENCE 757 AA; 86992 MW; 8079EB8BFA33EF21 CRC64;

Query Match 96.2%; Score 3687.5; DB 4; Length 757;
 Best Local Similarity 93.0%; Pred. No. 2.4e-268;
 Matches 707; Conservative 2; Mismatches 10; Indels 41; Gaps 3;

OY 1 MMSGVFYLLIFGKYFSGGGGQDVKCSIGYPCGNTKCLPQLLHCNGVDCCGNADSD 60
 |||||
 DB 1 MMSGVFYLLIFGKYFSGGGQDVKCSIGYPCGNTKCLPQLLHCNGVDCCGNADSD 60
 |||||
 OY 61 NC-----VVVLCQMSLPGLELDMKP- 82
 |||||
 DB 61 NCQDNGNMGQDFKPYASYKMTSQYFPEAEIPECLVGSVPVQCLOQ---GLELDCDEIN 117
 |||||
 OY 83 FTSVPSVSSNVTAMSIQWNLIRKLPPDCRKNYHDLOKLDLONNKITSISIAFRLNSLT 142
 :|||
 DB 118 LRAVPSVSSNVTAMSIQWNLIRKLPPDCRKNYHDLOKLDLONNKITSISIAFRLNSLT 177
 |||||
 OY 143 KYLSNRTTEPLKPGVEEDLHRLLEMLITEDNHLSTRISPPFGLNSLILLVLMNNVLTFL 202
 |||||
 DB 178 KYLSNRTTEPLKPGVEEDLHRLLEMLITEDNHLSTRISPPFGLNSLILLVLMNNVLTFL 237
 |||||
 OY 203 PKPPLCQHPRLHMLDLEGNHINHLNLTFTISCSNLTVLVMRKNKINHLENTFAPLOKI 262
 |||||
 DB 238 PKPPLCQHPRLHMLDLEGNHINHLNLTFTISCSNLTVLVMRKNKINHLENTFAPLOKI 297
 |||||
 OY 263 DELDLSGNTKIENLPILFDKLKLSQNLNLSYNPQIKIQANQDFIYLVKLSLSEGLEISN 322
 |||||
 DB 298 DELDLSGNTKIENLPILFDKLKLSQNLNLSYNPQIKIQANQDFIYLVKLSLSEGLEISN 357
 |||||
 OY 323 IOORMRPLMNLSHIYEFKFOYCYAPHYRSCPNFDGISLENLSTIQRFFVWVSA 382
 |||||
 DB 358 IOORMRPLMNLSHIYEFKFOYCYAPHYRSCPNFDGISLENLSTIQRFFVWVSA 417
 |||||
 OY 383 VYCFGNIFVYCMRPYIRSENKILYAMSIIISLCADCLMGITLVYIGGFDLKFGEYNKHAQ 442
 |||||
 DB 418 VYCFGNIFVYCMRPYIRSENKILYAMSIIISLCADCLMGITLVYIGGFDLKFGEYNKHAQ 477
 |||||
 OY 443 LMESTHCOLVGSLLSTFVSVLLTFLTEKYICIVPERCVRGKCTIVLLIWI 502
 |||||
 DB 478 LMESTHCOLVGSLLSTFVSVLLTFLTEKYICIVPERCVRGKCTIVLLIWI 537
 |||||
 OY 503 TGTIVAFIPLSNKEFFKNYGTNGVCFPLHSEPTESIGAOIYVAIFLGINLAFFIIVF 562
 |||||
 DB 538 TGTIVAFIPLSNKEFFKNYGTNGVCFPLHSEPTESIGAOIYVAIFLGINLAFFIIVF 597
 |||||
 OY 563 SYGSMFYVHQSATATEIRNOVKEMILAKRPFIVFTDALCWIPIFVVKFSLQVEI 622
 |||||
 DB 598 SYGSMFYVHQSATATEIRNOVKEMILAKRPFIVFTDALCWIPIFVVKFSLQVEI 657
 |||||
 OY 623 PGITTSWVVFILIPINSALNPILYTLTRPKEIHRFEMYNROKSMDSKQKTVAPSF 682
 |||||
 DB 658 PGITTSWVVFILIPINSALNPILYTLTRPKEIHRFEMYNROKSMDSKQKTVAPSF 717
 |||||
 OY 683 IWEEMWPLQEMPPELMKPDFTYPCESLSISOSTRLNSYS 722
 |||||
 DB 718 IWEEMWPLQEMPPELMKPDFTYPCESLSISOSTRLNSYS 757
 |||||
 RESULT 2
 O912Z5 PRELIMINARY; PRT; 737 AA.
 AC O912Z5;
 DT 01-DEC-2001 (Tremblrel, 19, Created)
 DT 01-DEC-2001 (Tremblrel, 19, Last sequence update)
 DT 01-DEC-2001 (Tremblrel, 19, Last annotation update)
 G PROTEIN COUPLED RECEPTOR AFFECTING TESTICULAR DESCENT.
 GN GREAT.
 OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RX MEDLINE=21250990; PubMed=11353515;
 RA Overbeek P.A., Gorlov I.P., Sutherland R.W., Houston J.B.,
 RA Harrison W.R., Boettger-Tong H.L., Bishop C.E., Agoulrik A.I.,
 RA "A transgenic insertion causing cryptorchidism in mice";
 RL Genesis 30:26-35(2001).
 RT EMBL: AF346501; AAL08943.1; .
 KW Receptor.
 SQ SEQUENCE 737 AA; 82943 MW; AF60F635EALACE49 CRC64;

Query Match 49.3%; Score 1891.5; DB 11; Length 737;
 Best Local Similarity 52.0%; Pred. No. 1.2e-133;
 Matches 370; Conservative 118; Mismatches 181; Indels 43; Gaps 7;

OY 27 CSLGTPCGNTKCLPQLLHCNGVDCCGNADSDNC-----VVVL-- 66
 |||||
 DB 28 CERKYPCCGNTKCLPRAHCDGVDCCGNADSDGDTSGWTTIFGYHGVNKKVTLIQ 87
 |||||
 OY 67 -----COCMSLPGLELDMKP-FTSVPSVSSNVTAMSIQWNLIRKLPPDCRKNYH 115
 |||||
 DB 88 ECPFLSYPOHCYCRE--NELECVKADLAVPKVSSNVTALLSLKKKHIRLPKAVFSRT 144
 |||||
 OY 116 DLQKLDLQNNKITSISIAFRLNSLTLYLSNRTTEPLKPGVEEDLHRLLEMLITEDNHL 175
 :|||
 DB 145 ELRKIYLQNNCTHISRRAFLGLHNLQIILYLSNCTISLPGIFKDLHQLAWLITDNDPI 204
 |||||
 OY 176 SRISPPFGLNSLILLVLMNNVLTFLPPDKPLCQHPRLHMLDLEGNHINHLNLTFTISC 235
 |||||
 DB 205 TRISQKSPFQNLNLFPLPWGNLEALPE-TLCAQPMQMLWDLNANGIKYITNSFTLC 263
 |||||
 OY 236 SNLTFLVMRKNKINHLENTFAPLOKIDELDGSNKIENLPILFDKLKLSQNLNLSYNP 295
 |||||
 DB 264 DSLITVFLPRNQGIFPERKFTSSLNKNGELDLSNNITKLPVLFSDHLQKLNLSNP 323
 |||||
 OY 296 IQRIQANQDFIYLVKLSLSEGLEISNIOORMRPLMNLSHIYEFKFOYCYAPHYRSC 355
 |||||
 DB 324 LLYVHKQFQSLKQLOSLDERIEIPNISTGMQPMKNLSHIYLFKFOYCYAPHYRSC 383
 |||||
 OY 356 PNNDGISLENLSTIQRFFVWVSAVYCFGNIFVYCMRPYIRSENKILYAMSIIISLCA 415
 |||||
 DB 384 PSTDGISSSEDLANGILRVSWVIAFITCVGNFLYAVASLIKAETHTAMSKILICA 443
 |||||
 OY 416 DCLMGITLVYIGGFDLKFGEYNKHAQLMESTHCOLVGSLLSTFVSVLLTFLTEK 475
 |||||
 DB 444 DCLMGITLVYIGGFDLKFGEYNKHAQLMESTHCOLVGSLLSTFVSVLLTFLTEK 503
 |||||
 OY 476 YCIVPERCVRGKCTIVLLIWIITGTVAFIPLSNKEFFKNYGTNGVCFPLHSEPTES 535
 |||||
 DB 504 FLIVYVPEFSSMLRGLKQTAVALASIIWVVGFLIAVPTFRDYDGNFGKNGVCEPLHYDQ 563
 |||||
 OY 536 TESTIGAOIYVAIFLGINLAFFIIVSYGSMFYVHQSATATEIRNOVKEMILAKR 595
 |||||
 DB 564 AEDPGSGYSLGIFLGNLAFLIYIVTMCSTIKITLQRAEVASHGKEVAVANRF 623
 |||||
 OY 596 FFIVFTDALCWIPIFVVKFSLQVEIIGTITSWVVFILIPINSALNPILYTLTRPKE 655
 |||||
 DB 624 FFIVFTDALCWIPIFVVKFSLQVEIIGTITSWVVFILIPINSALNPILYTLTRPKE 683
 |||||
 OY 656 MIHREWTNROKSMDSKQKTVAPSF IWEEMWPLQ-----EMPELMKP 700
 |||||
 DB 684 LKQDLHKHR-RKPIFKVKKRSLASIVWTDESSLKIALGVSIMKP 734
 |||||
 RESULT 3
 O9VBPO PRELIMINARY; PRT; 359 AA.
 AC O9VBPO;
 DT 01-DEC-2001 (Tremblrel, 19, Created)
 DT 01-DEC-2001 (Tremblrel, 19, Last sequence update)
 DT 01-DEC-2001 (Tremblrel, 19, Last annotation update)
 G PROTEIN COUPLED RECEPTOR AFFECTING TESTICULAR DESCENT.
 GN GREAT.
 OS Mus musculus (Mouse).

DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update).
DE 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE CS5042 PROTEIN.
GN CS5042.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Tracheata; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC SPRATIN-BERKELEY.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abell J.F., Agbayani A., An H.-U., Andrews-Plannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclob J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheefer F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissensbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL: AE003753; AAF56490.1; -;
DR FlyBase: FBgn0039354; CG5042.
DR InterPro: IPR002106; AA_trna_ligase_II.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCR_Rhodopsn.
DR PROSITE: PS00339; AA_trna_ligase_II_2; UNKNOWN_1.
DR PROSITE: PS00237; G_PROTEIN_RECPT_FL1; 1.
DR PROSITE: PS00262; G_PROTEIN_RECPT_FL2; 1.
DR G-protein coupled receptor; Glycoprotein; Transmembrane.
SQ SEQUENCE 359 AA; 40880 MW; 71913BEBD7C9E739 CRC64;

Query Match 15.6%; Score 599.5; DB 5; Length 359;
Best Local Similarity 39.9%; Pred. No. 4e-37;
Matches 132; Conservative 60; Mismatches 124; Indels 15; Gaps 7;

QY 349 PVRMCKPSTGVSFODLLSKPVLRYSAWVATLTIAGNVLVAGRPYRDENVAVTMV 408

Db 3 PVRMCKPSTGVSFODLLSKPVLRYSAWVATLTIAGNVLVAGRPYRDENVAVTMV 62
QY 409 IISCCADCLMGITLVYIGGFDLKFGRGYNNKHAQIMESTHCQIVGSAIISTEVSULL 468
Db 63 IRLNLADMLMGFLVLTGVODYRYRNEYKRVLDWITSMQCTLTGLTAVSSSEVSMIL 122
QY 469 TELFLKXICIVYPRCPGKCRIT-TVLIVITGPIVAFIPL-----SNKEPFKY 522
Db 123 AFMSIERLTIADPPRGHRSIGNRYMVALICMITGGLAVAPVLLRTSTLPYGSYS 182
QY 523 GTNGVCEPLHSEDTEISGAQIVSAIFGLINIAAFIIIVFSYGSMPYSVHOSATATEIR 582
Db 183 GT---CFPLIHIEAFPMG-WLYSAFVFLGNLLVLTAMLYTALLISWTRR-SATPL- 236
QY 583 NOVKKEMILARFFPIYTDLCWPIPIVAVFLSLQVEIGTITSWVITIPINSALN 642
Db 237 --TLIDCEFAVRFFPIYTDLCWPIIWMKIMWFENYNSIDDIYAMLVFVPLPLNSAVN 294
QY 643 PILVLTTRPPEKMIH-REWNYNRKRSMD 672
Db 295 PLVITFTTPKTRNQLFLGKKKITSRRKRAEA 325

RESULT 4
ID 095Y16 PRELIMINARY; PRT: 1012 AA.
AC 095Y16;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE GLYCOPROTEIN HORMONE RECEPTOR.
GN ACPHR.
OS Aserina pectinifera (Starfish).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Asterozoa;
OC Asteroidea; Valvatocoa; Echinodermata; Aserina.
OX NCBI_TaxID=7594;
RN [1]
RP SEQUENCE FROM N.A.
RA Mita M., Hirai T., Oba Y., Yoshikuni M., Nagahama Y.;
RT "cDNA cloning and functional analysis of a novel member of the
RT glycoprotein hormone receptor family from a starfish Aserina
RT pectinifera."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB061862; BAB68209.1; -;
KW Receptor.
SQ SEQUENCE 1012 AA; 112623 MW; 52A70E7A88C46E0A CRC64;

Query Match 15.0%; Score 576; DB 5; Length 1012;
Best Local Similarity 25.9%; Pred. No. 7.7e-35;
Matches 193; Conservative 102; Mismatches 262; Indels 188; Gaps 21;

QY 83 FTSYVS-----SNVAMSLOMNLIRKLPDCEKYNHDLQKLDLONNKITSISYAFR 136
Db 127 FQGYPRKAFRNDLANLKRHLDSMWIREVPADAMNLTALHLLDNLQSEVFTALH 186
QY 137 GUNSLFKLYSHNRITTPKGVF-EDLHRLFWLIEDNHLISPTPYGSLNLLVLM 195
Db 187 HLSNRIILHENSIPVVDHAFENSHLIE-LILRHKKITHLSAAVAGLPNWLLEFL 245
QY 196 NNVLRL-----PD-----KPLCQ 209
Db 246 GNSITSIHHTAFRNLPALRNVLILEVKNLSVPDLGTGTSLEHLIERCSLRAITANFCD 305
QY 210 HMPRLHWLDEGNHINLNLITFISCSNLTLYVMKKNINHLNENTFAPLOKIDELDGS 269
Db 306 NMTGLSLNLHNNLLEGPSLS-KCSSLKVHLCTNKLTSLEGQPFSGDLHYLDQLE 363
QY 270 NKIEMLPLIFKDLKELSQLNSYPYQIKQANQDYLVKLSLE-----GIE 319
Db 364 NDISYIPADAFOSLSHLTLTSLSNNTIREIDSQAFPCISLOYIDLNNSEFVPLTAGIQ 423
QY 320 I-----SNIQQRNFRPLMNLSHI-----YFKRFQYCGY----- 347

```

Db 424 MLKIRTYDNEQLEDFPPSELSITEIATAY--PYHCEYELEAEYKISLADRPNISE 481
OY 348 -----APHVSCKPN 357
Db 482 TTYWASGSVPDYNNMTFDINSESWIDSITFGSLSITGSPITYLSGNSRLLYVHNHNSCRK 541
OY 358 TDGISLENIASIIOREVVWVYSAVTCGNIYICMRPIYSENKLYAMSIISLCCADC 417
Db 542 PGPFMPCMDLFGSWPLRIGWLVFLAIGNAIVFVIIVSHTKMDVPREFLCLNLAFADE 601
OY 418 IMGITLVIGGDLKFRGYNKHAOLMSESTHCOLVGSALISTEVSVLLFLELEKI 477
Db 602 FLGVYLGFLAGVDTSTLGVFRKFGARWOLSAQCLAGFLAVSSEFSITLSTYITLERY 661
OY 478 CIYVPRCVPRGKC--RTITVLILMITGFIYVAFIPLSNKEFFKNYGTNGVCPPLHSEDT 536
Db 662 AIKHAHLHEKRKRLPHATVYMGCFWIFSVTAVALPLVN---VSHHRAVPCLPF--DV 714
OY 537 ESIGAOIYVAIFLGINLAFLIIVSYSGMFSYVHQS--AITATEIRNOYKKEMLAKRF 595
Db 715 DYTVAKVY--VGSILILNLAIFVIMACVASYIALIOGSHAMNCNDR-----VARRK 765
OY 596 FFIYFEDALCWPI-----FVYKELSLQVEIGTITSWVIFILPINSALNPLTY 647
Db 766 SLIVFTDFACWADIAPFSILTAAGFLRLISLDGAKV-----LTIYVPLNSCANPFIYT 818
OY 648 LITRPERK---EMIHFRWYN--YROR 667
Db 819 ILTKQFKKCKCKITMRSLSNRVFROR 843

```

```

RESULT 5
O95Y17 PRELIMINARY; PRT; 1280 AA.
AC 095Y17:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE GLYCOPROTEIN HORMONE RECEPTOR.
GN ACPGR.
OS Asterina pectinifera (Starfish).
OC Eukaryota; Metazoa; Echinodermata; Asterozoa;
OC Asteroidea; Valvatacea; Valvatida; Asterinidae; Asterina.
OX NCBI_Taxid=7594;
RN [1]
RP SEQUENCE FROM N.A.
RA Mita M., Hirai T., Oba Y., Yoshikuni M., Nagahama Y.;
RT "cDNA cloning and functional analysis of a novel member of the
RT glycoprotein hormone receptor family from a starfish Asterina
RT pectinifera."
RT Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AB061861; BAB68208.1; -.
KM Receptor.
SQ SEQUENCE 1280 AA; 141700 MW; 0AB0ECC0DD880BA CRC64;

```

```

Query Match 15.0%; Score 576; DB 5; Length 1280;
Best local Similarity 25.9%; Pred. No. 1e-34;
Matches 193; Conservative 102; Mismatches 262; Indels 188; Gaps 21;
OY 83 FTGVSPTS-----SNYATSLQWMLIRKLPDCEFKNYHDLQKIDLONKNTISITAFR 136
Db 127 FQGVPRKAFRNDLANLRKILHDSNMWIREVPADAFNMLTALHINLNDHNLSEVPTAALH 186
OY 137 GLASITFLYLSHRITFLKRGVF--EDHRLKEMLIIDNHLSTRISPPFYGNSLILVLM 195
Db 187 HLSNRLIHLHNSIPVDPHAFANSHLE--ILRHNKITHSAHAFAGLNPMLLEFL 245
OY 196 NVNLTFL-----PD-----KPLCQ 209
Db 246 GNSITSTIAHTAFRNLPLALRNVLLEVKNLVSVPDLTGTTSLHGLIGTERCSLRATPAFCD 305

```

```

OY 210 HMRPLHMLDLEGNHINLNFLEFISCSNLTVLMRKKNINHNENTFAFLQKLDELIGS 269
Db 306 NMTGLTSLNLHNNLIEGLPSLS--KCSILKYHLGTNKLITLSEGGPFGSLHLYDLQLE 363
OY 270 NKIENLPLPIFDKLELSQNLNYPNIOKIQANQFDYVLRKLSLE-----GTE 319
Db 364 NDISYIPADAFQSLHDLTSLSNNT IREIDQAFAPCTSLQYLDLSNNSFPVPLTAGIO 423
OY 320 I-----SNIOQRNRPPLMNSHT-----YFKFQYQGY----- 347
Db 424 MLKIRTYDNEQLEDFPPSELSITEIATAY--PYHCEYELEAEYKISLADRPNISE 481
OY 348 -----APHVSCKPN 357
Db 482 TTYWASGSVPDYNNMTFDINSESWIDSITFGSLSITGSPITYLSGNSRLLYVHNHNSCRK 541
OY 358 TDGISLENIASIIOREVVWVYSAVTCGNIYICMRPIYSENKLYAMSIISLCCADC 417
Db 542 PGPFMPCMDLFGSWPLRIGWLVFLAIGNAIVFVIIVSHTKMDVPREFLCLNLAFADE 601
OY 418 IMGITLVIGGDLKFRGYNKHAOLMSESTHCOLVGSALISTEVSVLLFLELEKI 477
Db 602 FLGVYLGFLAGVDTSTLGVFRKFGARWOLSAQCLAGFLAVSSEFSITLSTYITLERY 661
OY 478 CIYVPRCVPRGKC--RTITVLILMITGFIYVAFIPLSNKEFFKNYGTNGVCPPLHSEDT 536
Db 662 AIKHAHLHEKRKRLPHATVYMGCFWIFSVTAVALPLVN---VSHHRAVPCLPF--DV 714
OY 537 ESIGAOIYVAIFLGINLAFLIIVSYSGMFSYVHQS--AITATEIRNOYKKEMLAKRF 595
Db 715 DYTVAKVY--VGSILILNLAIFVIMACVASYIALIOGSHAMNCNDR-----VARRK 765
OY 596 FFIYFEDALCWPI-----FVYKELSLQVEIGTITSWVIFILPINSALNPLTY 647
Db 766 SLIVFTDFACWADIAPFSILTAAGFLRLISLDGAKV-----LTIYVPLNSCANPFIYT 818
OY 648 LITRPERK---EMIHFRWYN--YROR 667
Db 819 ILTKQFKKCKCKITMRSLSNRVFROR 843

```

```

RESULT 6
O9BN18 PRELIMINARY; PRT; 1050 AA.
AC 09BN18:
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE LEUCINE-RICH REPEAT-CONTAINING G PROTEIN-COUPLED RECEPTOR 2.
GN RK.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Nishi S., Hsu S.Y., Zell K., Hsueh A.J.;
RT "Characterization of two fly LGR (leucine-rich repeat-containing G
RT protein-coupled receptor) proteins homologous to vertebrate
RT glycoprotein hormone receptors: constitutively activation of wild type
RT fly LGR1 but not LGR2 in transfected mammalian cells."
RT Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF274591; AAK00808.1; -.
Db 127 FQGVPRKAFRNDLANLRKILHDSNMWIREVPADAFNMLTALHINLNDHNLSEVPTAALH 186
OY 137 GLASITFLYLSHRITFLKRGVF--EDHRLKEMLIIDNHLSTRISPPFYGNSLILVLM 195
Db 187 HLSNRLIHLHNSIPVDPHAFANSHLE--ILRHNKITHSAHAFAGLNPMLLEFL 245
OY 196 NVNLTFL-----PD-----KPLCQ 209
Db 246 GNSITSTIAHTAFRNLPLALRNVLLEVKNLVSVPDLTGTTSLHGLIGTERCSLRATPAFCD 305

```


QY 574 SATTAETIRNOYKEMILAKREFEIVFDALCMPIFVVKFSLLOVELPGITISWVIF 633
DB 569 PSSAAAHADTRV-----AQRMAVLFTPTDICAPISFPAISAAKLPL-LTVSDSKLL 621
QY 634 IL--PINSALNPILTYLTRPFR 654
DB 622 VLEFPINCSNPFLYAFETRNFR 644
RESULT 8
Q9NRK06 PRELIMINARY: PRT: 1300 AA.
Q9NRK06 Q9VJ03:
DR 01-OCT-2000 (Tremblrel. 15, Created)
DR 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE 01-JUN-2001 (Tremblrel. 17, Last annotation update)
GN HYPOTHEICAL.144.0 KDA PROTEIN (RK GENE PRODUCT).
OS RK OR BG:DS00180.13 OR CG8930.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Y, AND CN BW SP;
RX MEDLINE=9403001; PubMed=10471707;
RA Ashburner M., Mitra S., Roote J., Lewis S.E., Blazee R., Davis T.,
RA Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,
RA Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,
RA Palazolo M., Reese M.G., Spradling A., Tsang G., Wan K., Whitehead K.,
RA Celinker S., Rubin G.M.;
RT "An exploration of the sequence of a 2.9-Mb region of the genome of
RT Drosophila melanogaster: the Adh region.";
RL Genetics 153:179-219(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Y, AND CN BW SP;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Agapayni A., Arcaina T.T., Baxter E., Blazee R.G.,
RA Butenoff C., Champagne M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,
RA Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,
RA Houston K.A., Hummel S.R., Kaira K., Kearney L., Kim E., Lee B.,
RA Lewis S., Li P., Lomolan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,
RA Nixon K., Pacle J.M., Park S., Pfeiffer B., Poon L., Sequeira A.,
RA Sethi H., Snir E., Sylrskas R.R., Wan K.H., Weinburg T., Zhang R.,
RA Zieran L.L., Rubin G.M.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brannon R.C., Rogers Y.-H.C., Blazee R.G., Champagne M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Mills G.L.G.,
RA Abill J.F., Agapayni A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Boritova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Kocha S., Dunkov B.C., Dunn P.,
RA Durkin K.J., Evangelista C.C., Petraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Gloeck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegami C.,
RA Jalili M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Moberly C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Munzy D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nussbaum D.R., Pacle J.M.,
RA Palazolo M., Pitman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AEO03408; AAF4846.1; -;
DR EMBL: AEO03642; AAF53367.2; -;
DR FlyBase: FBgn0003255; tk.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR InterPro: IPR003592; LRR_out.
DR InterPro: IPR003591; LRR_typ.
DR Pfam: PF00001; 7tm.1; 1.
DR PRINTS: PR00237; GPCR_RHODPSN.
DR SMART: SM00370; LRR; 1.
DR SMART: SM00369; LRR_typ; 4.
DR PROSITE: PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 1300 AA; 144031 MW; B4B9E39F942FA0B3 CRC64;

Query Match 12.8%; Score 492.5; DB 5; Length 1300;
Best Local Similarity 22.6%; Pred. No. 1,9e-28;
Matches 187; Conservative 110; Mismatches 292; Indels 237; Gaps 21;
QY 45 LMCNGYDCGNQADEDCNCVYL-----CQMSLPGLE 76
DB 163 LSCRGIGIILAVPNLPNEVYVLTLSNDSIINMDPNAFYGLAKLRSLQNGCKSLSPPOS 222
QY 77 LDMKRFSTSPVSSN-----VTAMSLQNLIRKLPDDCKRYNDLOKLD 121
DB 223 FQGLAULTSL-QUNGALVSLDGDCHLOKLTIRKGNLFRTIPNALAGRTLEALL 281
QY 122 LQNNKITSIVYFRLNSLTLYLSHNRITFLKPGV-----FEDP-----HRLDW----- 167
DB 282 LKRNQMKISAGALKUTLAKVLELDNLISLPEGLSKLSQGLSTISNRRLWINDTE 341
QY 168 -----LTIEDNHSRISPPFFYGLNSLILVLMN-NVITRLPD----- 204
DB 342 LPRSMQMLMRANPLSTISPGARGMSKLRKLTISDVRLRSPPELEACHALEILKLDRA 401
QY 205 -----KPLCOHMRPLRHLMDLEGNHNLNLFIFSCSNLTIVYMKRNKINHLENTFAP 258
DB 402 GIOEVPANLCRQTPRLKSLDLSNQLKIOGKPFNGIKOLINDLISYNRKALPQAFQ 461
QY 259 LQKIDEL-----DLGSNKIEMLPPLFKDELKELSQLNSYNPQIKQANQPDYLV 308
DB 462 IPRQLQLYVSIIRIHSDELGNSIYHKAFSGFPALEDNLGNINPPELPSGLRALL 521
QY 309 KKLKSLSEGI-----EISNIQRMRP-- 330
DB 522 HKTFFNPKLRPPPPDFPRPQIQTILISYAHCACFLPVMSSQKTSQVOBAVLPSPD 581
QY 331 -----LMNL-----SHI-- 337
DB 582 AEFDMTLNNSNMNIMPQHNHNSKQLGASMDPWEALNFNEBQLOTQGGQATSYME 641
QY 338 YPKFQCYG-----YAPHVRSCKPNTDGSSLENTLASTIQHVFVW 378
DB 642 YFEHHVSGPAGYGGGTGLFSGMTEDPQPSGVOLPMPGFLPADLPDMWTYLRGFW 701
QY 379 VVSAYVTCFQGNIFVCMRPYRISRENKLYAMSISLCCADCLMGIFYLVIGFDLKRGEYN 438

Db	702	VFFLLSTLNGVFEVLLCSRSKMDVPREVLNCAADADEFMGVIGLILAYDAANTLGEFR	761
Oy	439	KHAQAMSTHCQOLVGSALISTEVSULLFLTEIKTICIVYPRCYRPPKCRITTVLI	498
Db	762	MEALPMQMSVSLCQSLGSLAVLSSELVYTLAVITLERVATTHAIHLKRLSLKQAGYIM	821
Oy	499	LI-WITGFEIVAFPLSNKEFEKNYGTGVCFFPLHSEDTESIGAOIYSVAIFLGINLAF	557
Db	822	SVGWFAIMLMLPVLGVSDYRKF-----ACVLPF--ETTGPASLIYIISLMF--INGCAF	874
Oy	558	IIVFSGYSMFYSVHQSAIATATEIRNOVKEMILAKREFEIVFDALCWIP-----I	609
Db	875	LTLMCYCLKMWALINGSOAWMT-----NDSRIKRMALLVETFDLWSPAFESITAI	927
Oy	610	FVVKRLSLQVEIPETITSWVIFLPIINSALNPILTYLTRPFE	655
Db	928	FGLOLITLSEQAKI-----FTVEVPLINSCNPFLYAIMTKQFKK	966
RESULT	9		
ID	0921P4	PRELIMINARY:	PRT: 907 AA.
AC	0921P4:		
DT	01-MAY-1999	(TREMBLrel. 10, Created)	
DT	01-MAY-1999	(TREMBLrel. 10, Last sequence update)	
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)	
DE	ORPHAN G PROTEIN-COUPLED RECEPTOR FEK.		
GN	GPCR49.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
CC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_Taxid-10090;		
RP	[1]		
RA	SEQUENCE FROM N.A.		
RA	MEDLINE-9912127; PubMed-9920770;		
RL	Herney G., Methner A., Schaller H.C., Hermans-Borgmeyer I.;		
RT	"Identification of a novel seven-transmembrane receptor with homology		
RT	to glycoprotein receptors and its expression in the adult and		
RT	developing mouse.";		
RL	Biochem. Biophys. Res. Commun. 254:273-279(1999).		
DR	EMBL; AF110818; AADI4684.1; -.		
DR	HSSP; P23945; 1XUN.		
DR	MGD: MGI:1341817; Gpr49		
DR	InterPro: IPR000276; GPCR_Rhodopsn.		
DR	InterPro: IPR001611; LRR.		
DR	InterPro: IPR000372; LRR_Nterm.		
DR	InterPro: IPR003592; LRR_out.		
DR	InterPro: IPR003591; LRR_typ.		
DR	Pfam; PF00001; 7tm.1; 1.		
DR	Pfam; PF00560; LRR; 15.		
DR	Pfam; PF01462; LRRNT; 1.		
DR	PRINTS; PR00037; GPCR_RHODOPSN.		
DR	PRINTS; PR00019; LEDRICHPRP.		
DR	SMART; SM00370; LRR; 3.		
DR	SMART; SM00013; LRRNT; 1.		
DR	SMART; SM00369; LRR_Typ; 8.		
DR	PROSITE; PS50262; G_PROTEIN_RECPE_FL2; 1.		
KW	Receptor.		
SO	SEQUENCE	907 AA; 99681 MW; 553167C6COAAE253 CRC64;	
	Query Match	12.8%; Score 490; DB 11; Length 907;	
	Best Local Similarity	24.6%; Pred. No. 1,9e-28;	
	Matches	182; Conservative 110; Mismatches 263; Indels 184; Gaps	24;
Oy	71	SILPGLDMKRFETISVPSSVSNVTASLQW--NLIRKLPPDCFFKNYHDLOKLDQNNKI	127
Db	139	SLOSRLDANHSIVPSCFSGIHSRLHMLDDNALTDVPVQAFPSLSLQAMTLANKI	198
Oy	128	TSISIIARGLNSLTKLYLSHRITFLKPGVFEDELHRLLEMLIIEENHLSRISPRFFYGIN	187
Db	199	HHIAIYAGNSSLIVLHLHNNRHISLCKKCFDGLASLETDLANNNLDEF-PTIRIKTLS	257

OY	188	SLILLYLNNVLTFRP-----DKPL-----COHPRPLHMLDLS-NH	223
Db	258	NKELGCHSNNSIRSIPEAFYGNPSLTIHFYDPIDFQVGSAAFOLEPLRTLTINASH	317
OY	224	IHNLRNLFEGCSNLTFLVWRKRNINHLNENTFAPLOKLDLDSGNKINLPLLPFKDL	283
Db	318	ITERPHILT-CTATLDESJTLTGANISSLPQAVCOQLPNLQYLDLSYMLDLDPSL--SGC	373
OY	284	KELSQLNLSYNPLOKIQANQFDIYVLLKSLSLBSIESLISNQQRFRPLMN-----	333
Db	374	OKLOKIDLRHNEIYEIKGSTFOOLFNRSLNLMANNKATAIHPNFAFSLPLSLIKLDLSNL	433
OY	334	-----LSHYF-----KKRQYCGY-----	347
Db	434	LSPFPVTGLHGLTHLKLGNBALOSLIPSANFPELKITEMPSAIVCCAFGECENYKISN	493
OY	348	-----APHRSCKPNFTDGISLEN	366
Db	494	QMNKDDGNSVYDDLHKKQAGLFQVQDERDLBEDFLDFEDLNALHSVGCSPSPGPKRCEH	553
OY	367	LLASTIQREVVWVAVTCEGNITV---ICMRPIYRSEKLYAMSIISLCCADCLMGITL	423
Db	554	LEGSMLIRIGVWTAVALTLCNALVAALTVEFTPLYSISIKL-LIGVIAV--VDILMGVSS	610
OY	424	FVIGSEFDLKFRGEVKNKHAQLMESTHOGLVSLALISTEVSALLFTLETKYICIIYPF	483
Db	611	AVLAADVFTTGERAQHGAHMEWDIGCOIVGFSLFASSESIPLTLAALERGSV---	666
OY	484	RCVRRGKC-----RTITVL-ILIMTIGFIAFPIPLSNKEFFKNYVGTNGVCF	529
Db	667	-----KCSKFEVKAPLFLSRALVILLCVILALF---IATPILG---GSKYVASPLCL	713
OY	530	PLHSDPTISGAQIYSVAFLGINLAAFIIVESYSGMFSYSVHOSATTAETIRQVAKEM	568
Db	714	PLPGEPEPTTG--YMAALV-LNSLCFLIMTIAVTLKLYCSLEKG-----ELEMWDCSM	764
OY	590	ILAKFEFVITDMLCWIPIVFRKFLSLQVE-IPGITSVWVFILPINSALPILYTL	648
Db	765	V--KHIALLEFRANCLICPVAFISFSSILNLTFSPPVIRFILIYVPLPSCLNPLXIY	822
OY	649	TYRPFKEMI-----HRFW	661
Db	823	FNPHEKEDMGSLGKHTRFW	841
RESULT	10		
O9ND11			
AC	O9ND11	PRELIMINARY;	PRT; 1360 AA.
DT	01-OCT-2000 (TREMBLrel. 15, Created)		
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
DE	GLYCOPROTEIN HORMONE RECEPTOR II.		
GN	RK OR BG:DS00180.13 OR CG8930.		
OS	Drosophila melanogaster (fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephydroidea; Drosophilidae; Drosophila.		
OX	NCBI_TaxID=7227;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-CANTON S.; TISSUE-WHOLE ANIMAL;		
RX	MEDLINE=20359836; PubMed=10899142;		
RA	Erikson K.K., Hauser F., Schiott M., Pedersen K.-M., Soendergaard L.,		
RA	Grimmelikhuizen C.J.P.;		
RT	"Molecular Cloning, Genomic Organization and Developmental Regulation,		
RT	and a Knock-out Mutant of a Novel Leu-Rich Repeats-Containing G		
RT	Protein-Coupled Receptor (DLGR-2) from Drosophila melanogaster.";		
RL	Genome Res. 10:924-938(2000)		
DR	EMBL; AF142343; AAF66608.1;-		
DR	HSSP; O57815; I03Y.		
DR	FlyBase; FBgn0003255; rK.		

```

DR InterPro: IPR000276; GPCR_Rhodopsn.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR003592; LRR_out.
DR InterPro: IPR003591; LRR_typ.
DR Pfam: PF00001; 7tm.1. 1.
DR Pfam: PF00560; LRR. 14.
DR PRINTS: PR00237; GPCR_RHODOPSIN.
DR SMART: SM00370; LRR. 2.
DR SMART: SM00369; LRR_TYP. 5.
DR PROSITE: PS0262; G_PROTEIN_RECEP_F1_2. 1.
DR Receptor.
SQ SEQUENCE 1360 AA; 150731 MW; 70435155B4F6F612 CRC64;

Query Match 12.7%; Score 486.5; DB 5; Length 1360;
Best Local Similarity 22.0%; Pred. No. 5,6e-28;
Matches 197; Conservative 106; Mismatches 277; Indels 315; Gaps 19;

DB 163 LSCRGIGILAVPVNLPRNEVYVLDGNNNLTREANSPFMAP-----NLEDLTLDSDN 213
OY 45 LHCNGVDCGNQADEDCVYVL-CQCMSLPGLELD--WMKPFSSVSVSSNVTAMSIQWN 101
DB 102 LIRKIPDCCKNTHDLOKLDLQNNKITSISITAFRGINSITKLYLSNRTTEFLKPGVFED 161
DB 214 SIINMPNNAEYGLAKLRSLQNCGLKSLPPQFOGLQTLSDLNGLNALVSLDGCLGH 273
OY 162 LHRLEMLIIDNHLRSIPTEFYGLNSLLVLMNNVLTPLDPKPLGQHPRLHMLDLEG 221
DB 274 LQKLTIRLQGNLFYRPTNALAGLRLEALNGLSNLLTIINDEDF-PRAPNLIVLLKKR 332
OY 222 NHIN-----LRNLTFISC-----S 236
DB 333 NQIMKISAGALKNLTKLVLELDNLISSLPEGLSKLSQGLSTISNRLRWINDTELPR 392
OY 237 NLIVLVARRKKINHLNNTFAPLQKDE----- 264
DB 393 SKMDLMDRAPLSTISAGARGMSKLRLISDVTRLRSEPELEACHLEILKLDRAQIQ 452
OY 265 -----LDGSKNIENLPLIFKDKELS 287
DB 453 EVRANLCROTPLRLKSLKLTNSLKRIPNLSSCRDLRLDLSNDETKIOGKPFNGLKQLN 512
OY 288 QLNLSYPIQIQANQDFYLVKLSLSEGLEISNIQRMF----- 328
DB 513 DLLSTYRIKALPQDAFOGIPKIQLDLENEISYIHKFAFSGFALDNLGNINPEL 572
OY 329 -----RLMLNSHIYFK-----FOYGYAHVAVSCK----- 355
DB 573 PESGLRLHLKTFNNKLEPPPTFPRIQTLILSYAVHCAFLPLVAMSSOKTSQY 632
OY 356 -----PNTDGIS----- 362
DB 633 QEAVLPESDAEFDMTLMNNSMMNITWPMNHLNSQLGASMDPMTATINFEQLOQTGG 692
OY 363 -----SLENLASTIQ----- 373
DB 693 QIATSYMEEFEEHDVSGPATGYGFGTGLFSGSTEDFOGVOCLPMPGPLPCADLFD 752
OY 374 -----RPFVWVAVATCGNIFVICMRPTSEKNTLYAMSIISLCCADLMGIYLFVIGGF 429
DB 753 WMTLRGVWVVFLLSLGNGTVFVLICSRSKMDVPFVLICNLAADFEGIYGLIAIV 812
OY 430 DLKFRGEYNHAOLMSESTQCLVGLAISTEVSLLLFLEFKYICIVYPPRCVRPG 489
DB 813 DATLGEFRFAIPQMSYVQLSGFLAVSSELYTTLAVITLERVATTHAIHLNKL 872
OY 490 KCRITVLLII-WITGFIAPLSNKEFFKNYTINGVCFPLHSEDTESIGAQISVAI 548
DB 873 SEKOAGIIMSVGWVFLIMALMPLVGSYDRKF---AVCLPF---ETTTGPASTIYVSL 926
OY 549 FLGINLAALFIIFYSGSMYSVHOSAITATELRNQYKKEMILAKRFFIVFDALCWP 608
DB 927 MF-INGCAFLTLMGCIYKMTWALIRGSOAMNT-----NDSRIAKRMALLVFTDFLCWSP 978

```

```

OY 609 -----IFVAFELSLQVEIDPQTTSWVIFILPINSALNPLYTLTPRPKE 655
DB 979 IAFSTIAIFGLDLSIEQAKI-----FTVFPLNSCNPFLXAIMTKQPKK 1026

RESULT 11
O9PVN9 PRELIMINARY; PRT; 658 AA.
ID O9PVN9;
AC O9PVN9;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE GONADOTROPIN RECEPTOR I.
GN sc9h-RI.
OS Oncorhynchus rhodurus (amago).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=41164;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20025349; PubMed=10558873;
RA Oda Y., Hirai T., Yoshitura Y., Yoshikuni M., Kawachi H., Nagahama Y.;
RT "The duality of fish gonadotropin receptors: cloning and functional
RT characterization of a second gonadotropin receptor cDNA expressed in
RT the ovary and testis of amago salmon (Oncorhynchus rhodurus).";
RL Biochem Biophys. Res. Commun. 265:366-371(1999).
DR EMBL: AB030012; BAA8698.1; -.
DR HSSP: P22888; 1LUT.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR InterPro: IPR001611; LRR.
DR Pfam: PF00001; 7tm.1. 1.
DR Pfam: PF00560; LRR. 3.
DR PRINTS: PR00237; GPCR_RHODOPSIN.
DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE: PS0262; G_PROTEIN_RECEP_F1_2; 1.
DR Receptor.
KW Receptor.
SQ SEQUENCE 658 AA; 72991 MW; 853A38BEA845480 CRC64;

Query Match 12.6%; Score 483.5; DB 13; Length 658;
Best Local Similarity 24.1%; Pred. No. 4e-28;
Matches 171; Conservative 113; Mismatches 274; Indels 151; Gaps 26;

OY 64 VVIC--CQMSLPGLELDWM-----KPFSSVPS-VSSNVTAMSIQWNLIRKLP 108
DB 10 MLICVLGCVSMQAEVAMVNSGTFYIYLCMGNTITMPTIIPKNTIDLERKQTHIRVFPQ 69
OY 109 DCFKNYHDLOKLDL-QNNKITSISIAFRGLNSLTJKLYLSHNRITFLKPGVFEDLHRLFW 167
DB 70 EAFNTNQQLALAIYLTENGMLSEIGAFANLPRLTET----- 106
OY 168 LIIEDNHLRSIPPTFYGLNSLLVLMNNVLTPLDPKPLCQIMPRH-----WIDLEG 221
DB 107 TIRKSKLYIIHQAFGLRKLKSLHLYICNTGLRVLP-----NFSRIHSAAMFELDLQD 160
OY 222 N-HIHNLRNLTFTSCSLTV--LVMRKN-----KINHLNEN 254
DB 161 NHIIVIPSNAPFGLTNTIDELRLTKNGISEVESHAFNGTKIHKLYLMGNLDLSHMHN 220
OY 255 TFAPLOKLELDGSKNIENLPLIFKDKELS-----QLNLSYN 294
DB 221 SFKGAEGPGLDISRALSSLPESVAGEVHLSAVSVSLRTLPLSLFTKLRQANLTPY 280
OY 295 -----PIQIQANQDFYLVKLSLSEGLEISNIQRMFR-LMNLSHIYFKFOYGYAP 349
DB 281 SHCAFHKKHQRN-----TFRMTSACFKPGQANNLHFF---MDFLNWT 321
OY 350 HVKSCPNNDGISLNLASTIQRFVWVAVATCGNIFVICMPTISEKNTLYAMSI 409
DB 322 SV-ACSBAPDAFPCEDIMGSAFLRVILITISVLALIGNTIVLLVLLGSAKMTVERFLM 380

```

```
QY 410 ISLCADCLMGIFYLVIGFDLKRGEYKNAQLMESTHCQVLSLAISTEVSLLLT 469
      || ||||| : : : || : : : || : : : || : : : || : : : ||
Db 381 CHSLFALCAGIYIVYATADYRTGRGLYNHAIQMGACDINGFTVPASELSMFTLT 440
QY 470 FLELEKTYICIVYPRCVPRPKCR-TITVLILWITGFIYAFILSNKEFFKNYGTNGVC 528
      :|||: : : : : : : : : : : : : : : : : : : : : : : :
Db 441 ALTLEHRTHTHALRLDKRLRLHACAVMAMGAFSCIALMLLPVGVG---VSSYSKVSIC 496
QY 529 FPLHSEPTESIGAQIYVAIFLGINLAIFIIVSYSGSMFYSVHQSATATATEIRNQKKE 588
      || : : : : : : : : : : : : : : : : : : : : : : : :
Db 497 LPM---VESLPSQVF-VMLLLNLVAFGLVCVCYLSTIYSVRNSS-----SPPA5AE 546
QY 589 MIARKEFFIYFDALCMIPFVYVKFLSLQVEIPGTTSMVVFIL--PINSALNPLIY 646
      :||: :|||: || : : : : : : : : : : : : : : : : : :
Db 547 TCMQGRRAIILITFDLCMAPISFPALSAALKPLI-TVSDSKLLLVLFYPIINSCANFLY 605
QY 647 TLTPRPKEIMHRFWYVRQKSM-DSKGQKYAPSF-----IIVEMKP 689
      ||||| : : : : : : : : : : : : : : : : : : : : : :
Db 606 GLCTRPR---RDFLLAARYGLFTTKAQYRTSEFSVOQAAMIQKSP 650
```

RESULT 12

```
Q9DGC5 PRELIMINARY; PRT; 693 AA.
AC Q9DGC5;
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE GONADOTROPIN RECEPTOR II.
GN TGTB-RII.
OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Labroidae;
OC Cichlidae; Oreochromis.
NCBI_TaxID=8128;
RN [1]
RP SEQUENCE FROM N.A.
RA Oba Y., Hirai T., Yoshitura Y., Yao Z., Nagahama Y.;
RT Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL: AB041763; BABI6107.1; -.
DR HSSP: P22888; ILUT.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR InterPro: IPR001611; LRR.
DR Pfam: PF00001; 7tm_1; 1.
DR Pfam: PF00560; LRR_2.
DR PRINTS: PR00237; GPCRHHODPSN.
DR PROSITE: PS00237; G_PROTEIN_RECP_F1_1; UNKNOWN_1.
DR PROSITE: PS50262; G_PROTEIN_RECP_F1_2; 1.
KM Receptor.
SQ SEQUENCE 693 AA; 77856 MW; C3BBFDEBCEFC8988 CRC64;
```

Query Match 12.5%; Score 481; DB 13; Length 693;
Best Local Similarity 25.4%; Pred. No. 6.6e-28;
Matches 156; Conservative 118; Mismatches 235; Indels 104; Gaps 24;

```
QY 103 IRKPPDCEKYNHDLQKIDL-QNNKITSISYIYARGLSLTKLYLH-NRTITLPGVVE 160
      || : : : : : : : : : : : : : : : : : : : : : : : :
Db 64 LKRPVSAFRELIIITIEISQSDCIHQHAFLSYLSAQISVONINSIRLFEGFA 123
QY 161 DLHLEWLIIEDNLHSRPTPEYGLNSLLLVLMNNVLPFLPKPDLCOHMRILHM---- 216
      || :|||: : : : : : : : : : : : : : : : : : : : : :
Db 124 DLPLLETLISNIGIAHF--PDFTTISS-----LSPNILEMANMEIDITIPANSFGCIT 176
QY 217 -----LDLEGNHINHLNLFISCSNLTVLVYMRKN-KINHLNENTFAPLQKIDELDGSN 270
      :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 177 EEEYDMNLVNRGEFEIKSHAF-NGTKLNTLVLRDNNWYLRNIGQDAFEGAGPTLLDVST 235
QY 271 KIENLPPLIFDKELKLSQNLNINYPKIQANQPDYLVKIKLSL-----EGIF----- 319
      : : : : : : : : : : : : : : : : : : : : : : : :
Db 236 ALRLSP-----NGLRHVYFLKASHAVALKSLPLLSLAELEALTYLP 279
```

```
QY 320 -----ISNIQQRMRPLN-----LSHIYFKKQY-----CGY 347
      :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 280 SHCCAFHTWRRKROKESALKNLTK--FCDLNTEIDTPADDTSLNDINFDLEPDCFS 337
QY 348 APHVRSCKPMTDGISISLENLAIIOQVFWVVSAYVCFGNIFICMRPYIRSENKLYAM 407
      || :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 338 NPEYK-CSPKPDANPCEDDLGFSFLRCLTWIINVEFVAGNLANVLI--LLIGHNKLTVS 394
QY 408 SIT--SLCCADCLMGIFYLVIGFDLKRGEYKNAQLMESTHCQVLSLAISTEVS 465
      : : ||||| : : : : : : : : : : : : : : : : : : :
Db 395 RFLMCNLAFADLCGLYLLIAFMDYSHHYYNHADWOTGPGCGIAGELYVSSLSV 454
QY 466 LLLEFLLEKTYICIVYPRCVPRPKCR-TITVLILWITGFIYAFILSNKEFFKNYGT 524
      ||||| : : : : : : : : : : : : : : : : : : : : : :
Db 455 YTLTVISLERHNTTANAHYKRLRMLHHVYAMVGNVAFSLVALDPLVG---VSSISK 510
QY 525 NGVCFPLHSEPTESIGAQIYVAIFLGINLAIFIIVSYSGSMFYSVHQSATATATEIRNQ 584
      :||: :|||: || : : : : : : : : : : : : : : : : : :
Db 511 VSICLPM---DIDPLGAQVYVAVLI-LNVYAFGLVCVCYLSTIYSVRNSS--RGD 564
QY 585 VKKEMIAKREFEIVFDALCMIPFVYVKFLSLQVEIPGTTSMVVFIL--PINSALN 642
      || :||: :|||: || : : : : : : : : : : : : : : : : :
Db 565 TK---IAKRAVLIIFFDLCMAPISFPALSAALKPLI-TVSHSKILLIFYPINSLCN 619
QY 643 PIYTLTPRPKE 655
      ||||| :||: :||: :||: :||: :||: :||: :||: :||:
Db 620 PLYTITTPRAFRK 632
```

RESULT 13

```
Q9DGF5 PRELIMINARY; PRT; 696 AA.
AC Q9DGF5;
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-DEC-2001 (TEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE FOLLICLE-STIMULATING HORMONE RECEPTOR PRECURSOR.
OS Cynops pyrrhogaster (Japanese common newt).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandridae; Cynops.
NCBI_TaxID=8330;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA MEDLINE=20403884; Pubmed=10944452;
RA Nakayama Y., Yamamoto T., Oba Y., Nagahama Y., Abe S.-I.;
RT "Molecular cloning, functional characterization, and gene expression
RT of a follicle-stimulating hormone receptor in the testis of newt
RT Cynops pyrrhogaster."
DB Biochem Biophys. Res. Commun. 275:121-128(2000).
RL EMBL: AB005587; BABI3501.1; -.
DR HSSP: P23945; IXUN.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000372; LRR_Nterm.
DR Pfam: PF00001; 7tm_1; 1.
DR Pfam: PF00560; LRR_4.
DR PRINTS: PR00237; GPCRHHODPSN.
DR SMART: SM00013; LRRNT; 1.
DR PROSITE: PS00237; G_PROTEIN_RECP_F1_1; UNKNOWN_1.
DR PROSITE: PS50262; G_PROTEIN_RECP_F1_2; 1.
KM Signal; Receptor.
FT SIGNAL 17
FT CHAIN 18 696 POTENTIAL.
SQ SEQUENCE 696 AA; 78633 MW; 179A6FC800B71E57 CRC64;
```

Query Match 12.5%; Score 479.5; DB 13; Length 696;
Best Local Similarity 23.7%; Pred. No. 8.6e-28;
Matches 165; Conservative 133; Mismatches 250; Indels 149; Gaps 30;

```
QY 32 FPGCNITKCLFQLLHNGVDGCGNQADEDCNVVVLCCQMSLPGLDMKMPFTSVP-SVS 90
```

```

Db      16  FCGHPVCRCLNRVETC-----QESHVY-----QIPRDIP 44
QY      91  SNTAMSLOMNLIRKLPPDCKNYHDQKDL-QNNKITSIYAFRGLNLTLYLSH- 148
Db      45  RNSLELFVLTQVIVIPKAASGEDEVENIESONDLAKTEANVFHLLKRLRIEKA 104
QY      149  NRITFLKPGVEDLHRLMWLIIEDNHLRSISPPFYGINSLLLVANNVLTRLPDLPC 208
Db      105  NNLVYIDPDAFQNPISLKYLLISTNGIQV--PAVSKIRS-----PHSVL----- 147
QY      209  QHMRPLMWLDEGN-HIHNLRNLTFISCSNLVLYV-MRKKINHLENFPAFLQKDELD 266
Db      148  -----LIVODNINIRHIGKNSFAGLSSESTIRLNKNGIEIQHAFNGTH-LNELN 198
QY      267  LGSN-KIENLPLIFKD-----LKEISQMLNLYNP-----IQIQANOQDYLVKLSLS 314
Db      199  LSDNQRLKLPDOVFQAGATPVILDISRTRIHFLPNNGIENIKFRARFNYLKLPPLE 258
QY      315  --LEGIE-----ISNIOQM--FRPLMN-----L 334
Db      259  KPELLEIANLTYSHCCAFANRERKSEMPICNKSFGKHSACKPEDKNLRRESNEDYL 318
QY      335  SHIYFK-----KEYOY---CGYADHVRSCKPNIDGISSLENLASIIQRYVFWVVA 382
Db      319  SSYGFSTSLVENGDEPNYDYLCEV--HDVYCEPKPDAPNCEIDIMQDNTLRVLIWLSI 377
QY      383  VTQGNFVFCMRPIYSEKNIKVAMSI--SLCCADLMGITYLFVIGPDLKFRGEVKNH 440
Db      378  LAITGNITVAVI--LISSQYKFTVPFPLMCLNAPADLCMGITYLLIASVQIKRKSQYVNH 435
QY      441  AOLMESTHCQVLSLALISTEVSALLTFLETKYICYVFCRCVPRGCRITVYILI 500
Db      436  AIDMQTSGCAAGFEVFAFVSELSVLYLTLEMRHTITYAMQDLKVRFRHATAIMV 495
QY      501  -WITGFVAITPLSNKEFFKNYGTNGVCPPLHSEDTESIGAQIYVAIFLGINLAFTI 559
Db      496  GWIEFAFVALPI---FGVSSYTKVASICLPM---DIESRISOSIYFI-LVLMVCAFLI 547
QY      560  IVSYSGSMFVQHSATTAITEIRNOVKEMILAKRFFIYFDALCWIPIFVVFLELQ 619
Db      548  ICACYIGIYLVTRNPNVSS-----NSDKIKAKRAIILFTDPLCANPISFAISLSLK 601
QY      620  VEIPGTTTWNVFIIL--PINSALNPILYLTTRPFK 654
Db      602  IPII-TVSKSKILLVLFYPINSCANPFLVYAIFFKTFR 637

RESULT 14
Q9PM16  PRELIMINARY; PRT; 662 AA.
AC 09PM16;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, last sequence update)
DE 01-JUN-2001 (Tremblrel. 17, last annotation update)
DE FOLLICLE-STIMULATING HORMONE RECEPTOR PRECURSOR.
GN FSH-R.
OS Clarias gariepinus (Sharpooth catfish) (African catfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Siluriformes; Clariidae; Clarias.
OX NCBL_TaxID=13013;
RN [1]
RN SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA Bogerd J., Andersson E., Blomenroth M., Tensen C.P., Gramanman J.C.,
RT Schultz R.W., Geos H.U.;
RT Cloning and functional characterization of a testicular follicle-
RT stimulating hormone receptor of the African catfish.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; A0102647; CAB51907.1; -.
DR HSSP; P23945; 1XUN.
DR InterPro; IPR000276; GPCR_Rhodopsn.

```

```

DR InterPro; IPR001611; LRR.
DR Pfam; PF000601; 7tm_1; 1.
DR Pfam; PF00560; LRR; 4.
DR PRINTS; PR00237; GPCRHOPOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW Signal; Receptor.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 662 FOLLICLE-STIMULATING HORMONE RECEPTOR.
SQ SEQUENCE 662 AA; 74891 MW; 5CB49EDDBD1DAF5 CRC64;

Query Match 12.5%; Score 478; DB 13; Length 662;
Best Local Similarity 25.4%; Pred. No. 1,le-27;
Matches 157; Conservative 94; Mismatches 257; Indels 110; Gaps 21;

QY 92 NVTAMSLOMNLIRKLPPDCKNYHDQKDL-QNNKITSIYAFRGLNLTLYLSHNR 150
Db 49 NTVYVEIKLQIIFPRAMSSLDLKRIMVSENGALQRIEAVAFANLTLEEI----- 102
QY 151 ITFLKPGVEDLHRLMWLIIEDNHLRSISPPFYGINSLLLVANNVLTRLPDLCOH 210
Db 103 -----TITKSKNLVSMRDTFGPLKRLITISNTGLVLPDPSKYQS 145
QY 211 MPRLHWLDEGN-HIHNLRNLTF--ISCSNLVLYVMRKNINHLENFTF----- 256
Db 146 AAEFLPDLDDNMHIEVTHSNAFAGLISGTTTELRLTKNGITEVERNAFNGTKMEKFLM 205
QY 257 --APLOKDE-----LDGSKNIENLPLIFKDK----- 264
Db 206 GNOOLKRIDNHAFLEAGPLVLDISRTAISLSPENMRIRKLLIANSVLSRLVPLNLEIF 265
QY 285 -ELSQLMSTNPNLOKIQANOQDYLVKLSLLEGIEISNIOQRMFRPLMLSHIYFKFQ 343
Db 266 AELTQANLTY-PSHCAFKNF-----KKNLEKNHLCNVS-----TIRQDEYFFP-E 311
QY 344 YCGYAPHRSCKPNIDGISSLENLASIIQRYVFWVVAVTCFNGNIFVICMRPIYSEK 403
Db 312 HCKRVLEV-SCYEPDPAFNPCEIDIMGFTFLRWLIMFISVLAUNGPFVLLV--LLSHTK 368
QY 404 LYAMSI--SLCCADLMGITYLFVIGFDLKRGEVKNHQAOLMESTHCQVLSLALIST 461
Db 369 LTVPKFLMCHLAFADFCMGILILIASVDLTQSRYINBIEQOTGCGTAGFLYFAS 428
QY 462 EVSALLTFLETKYICYVFCRCVPRGCR---TIVLLIIMITGFIYVFIPLSNKEFF 518
Db 429 ELSVYTLTATLEBRMHTITYAMRLER--ECRLNHACCVMVAFGWVSYLAALMPYIG---- 482
QY 519 KNYGTNGVCPPLHSEDTESIGAQIYVAIFLGINLAFTIYFSGSMFVSHQSATA 578
Db 483 VSSYMKTSICLPM--DVEIVSSQVY-VMLLEFLNVLAFVAVSCYVRIYITVRHPASVP 538
QY 579 TEIRNOVKKEMILAKRFFIYFDALCWIPI-FVVKRFLSLQVEIPGTISWVYFIPLI 637
Db 539 D-----SADARAKKRAVAVFTDFLCMAPISEFAISALRQPLITYSHAKVLLVLFPI 592
QY 638 NSALNPILYLTTRPFKE 655
Db 593 NSCANPFLVYAIFFKSPFQ 610

RESULT 15
Q9YVG0  PRELIMINARY; PRT; 334 AA.
AC 09YVG0;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, last annotation update)
DE CG4187 PROTEIN.
GN CG4187.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

```

OX Ephydroidea: Drosophilidae: Drosophila.
 NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=7131132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Maniatis P.G., Scherer S.E., Li P.W., Hoskins R.A., Gale R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazei R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agayuni A., An H.-J., Andrews-Plannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Bendale J., Bayraktiroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
 RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Gang N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Correll J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwan C.,
 RA Jalali M., Kalish F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merklow G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reaenert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Tiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Slier E., Spreading A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassatman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003491; AAF48237.1; -
 DR FLYbase: FBgn030458; CG4187.
 DR InterPro: IPR000276; GPCt_Rhoopsn.
 DR InterPro: IPR001901; SecE.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PR00237; GPCRHHOOPSN.
 DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
 DR PROSITE: PS01067; SECE_SEQ61G; UNKNOWN_1.
 QO SEQUENCE 334 AA; 37020 MW; 1652F2D467534061 CRC64;

Query Match	12.5%	Score 477.5;	DB 5;	Length 334;
Best Local Similarity	36.1%	Pred. No. 5,2e-28;		
Matches	103;	Conservative	59;	Mismatches 114; Indels 9; Gaps 5;

QY	380	VSATYCEGNITVICMRPIRSENKL-YAMSIISLCCADCLMGITFYVIGEDPKRGEYN	438
	
Db	1	MASIAVGNLNLVIGRYFYKRSRNSVHSLYLRHIAASFLMGIIYTLTLACADISFREGDI	60
	
QY	439	KHAOLMESTHCOLVGLAISTEVSVLLEFLTEKYICIVYPPRCVAPGKCFITVLI	498
	
Db	61	KYEETWRHSGCAFGAFSTFSCOSTLTLLTVMDRLMSVTRPQKPRDEKVIYLRLL	120
	
QY	499	LIMITGFVATLP-LSNKEFPKNYGTGNGVCPHLSSEPTESGAOITYSAIRGLINIAAF	557
	
Db	121	LIMGSEFLAAAPLIPNYPFGSHFTGNNGVCLSLIHDPYAGWE-TSALTELTIVNNTSL	179
	
QY	558	TIIVFSYGMFYSHQOS--AITATEIRNOVKKEMILARFPEIVETDCLCWIPIFVKPL	615
	

```

D6 180 IFLESTIRMLQALRDSGGGRST----HSRENVVATRAILYTTDCAOMLPIYVKA 235
Q7 616 SLQVEITPGTTSWVVFILPINSALNPILYTLTTRPKEMIRF 660
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
236 ALSCEISPDLYAMLVAVLPVNSALNPVLYTLTTLTAFFKQQLRY 280

```

Search completed: September 5, 2002, 09:02:58
Job time: 2302 sec

